Running head: Coding Accuracy of the Ambulatory Data System

Coding Accuracy of the Ambulatory Data System:

A Study of Coding Accuracy

Within the General Internal Medicine Clinic,

Walter Reed Army Medical Center

A Graduate Management Project

Submitted to the Faculty of Baylor University

In Partial Fulfillment of the Requirements for the Degree

Of

Master of Health Care Administration

By

MAJ Daniel W. Gall, MS, USA

17 April 1998

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE 17 April 1998	3. REPORT TYPE AND DATES COVERED FINAL REPORT (07-97 TO 07-98)

4. TITLE AND SUBTITLE
CODING ACCURACY OF THE AMBULATORY DATA SYSTEM: A STUDY OF DATA ACCURACY IN THE GENERAL INTERNAL MEDICINE CLINIC, WALTER REED ARMY MEDICAL CENTER

5. FUNDING NUMBERS

6. AUTHOR(S)
MAJOR DANIEL W. GALL, MEDICAL SERVICE CORPS, U.S. ARMY

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
WALTER REED ARMY MEDICAL CENTER
6825 16TH STREET N.W.
WASHINGTON, D.C. 20307-5001

8. PERFORMING ORGANIZATION REPORT NUMBER
9-98

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US ARMY MEDICAL DEPARTMENT CENTER AND SCHOOL
BLDG 2841 MCCS-HRA US ARMY-BAYLOR PROGRAM IN HCA
3151 SCOTT RD SUITE 1412
FORT SAM HOUSTON TEXAS 78234-6135

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED

12a. DISTRIBUTION / AVAILABILITY STATEMENT

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The Military Health System (MHS) has implemented a relatively new automated information system to help capture diagnoses, procedures, and insurance data for all ambulatory encounters. This system, recently implemented at Walter Reed Army Medical Center (WRAMC), is called the Ambulatory Data System (ADS). While the current MHS metric for ADS focuses on compliance, the quality of the data has yet to be extensively measured. Hence, the purpose of this project was to statistically analyze the data quality by studying the coding accuracy of International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes, Evaluation and Management (E&M) codes, and insurance indicator codes within the General Internal Medicine Clinic (GIMC) at WRAMC. Interventions to improve the data quality were identified, developed, and implemented during the course of this project. The sensitivities, positive predictive values (PPVs), and the kappa statistics determined from random samples collected before and after the interventions were compared using Chi square analysis (alpha = 05). The results showed an increase in overall non-adjusted ICD-9-CM coding rates from 60% to 67% sensitivity, 66% to 73% PPV, and kappa = .18 to .36, however, the differences were not significant. E&M coding improved from a poor sensitivity rate of 21% to significantly better rate of 55%. The study also identified a poor level of accuracy concerning the capture of insurance information (kappas = .18 and .16) that conservatively indicated approximately $1.35 million dollars of missed third party collections within the past year. This study provides a model to improve the quality ADS data that may enhance the organization's ability to efficiently and effectively identify clinical process improvements, make sound resource allocation decisions, increase third party collections, and conduct outcomes research.

14. SUBJECT TERMS
AMBULATORY DATA SYSTEM; CODING ACCURACY

15. NUMBER OF PAGES
73

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
N/A	N/A	N/A	UL

Acknowledgments

This project is submitted with great thanks and appreciation to several key individuals who, without their help, this project would not have been possible. I would first like to thank LTC David L. Jones, M.D., Chief, General Internal Medicine Service, for giving full support to this project by subjecting the General Internal Medicine Clinic to evaluation, and offering his invaluable assistance, insight, and time to participate in this project. Many thanks to Joyce Jones, A.R.T, DRG Coordinator, WRAMC, for volunteering her time, energy, and expertise during the many long hours involved in re-abstracting medical records, developing and conducting provider coding education presentations, and being readily available for consultation on coding matters. I'd also like to thank COL Jill Phillips, Chief, Utilization Management, WRAMC, for prompting the need and assisting with support of this project. Special thanks to my preceptor, COL Robert J. Heckert, Jr., Deputy Commander for Administration, WRAMC, for his support, patience, mentorship, and leadership.

Last, I'd like to thank God for blessing me with two beautiful children, Lauren and Patrick, and a truly wonderful wife, Christine, whose never-ending patience, love, and energy have been my guiding light throughout this project.

Abstract

The Military Health System (MHS) has implemented a relatively new automated information system to help capture diagnoses, procedures, and insurance data for all ambulatory encounters. This system, recently implemented at Walter Reed Army Medical Center (WRAMC), is called the Ambulatory Data System (ADS). While the current MHS metric for ADS focuses on compliance, the quality of the data has yet to be extensively measured. Hence, the purpose of this project was to statistically analyze the data quality by studying the coding accuracy of International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes, Evaluation and Management (E&M) codes, and insurance indicator codes within the General Internal Medicine Clinic (GIMC) at WRAMC. Interventions to improve the data quality were identified, developed, and implemented during the course of this project. The sensitivities, positive predictive values (PPVs), and the kappa statistics determined from random samples collected before and after the interventions were compared using Chi square analysis (alpha = 05). The results showed an increase in overall non-adjusted ICD-9-CM coding rates from 60% to 67% sensitivity, 66% to 73% PPV, and kappa = .18 to .36, however, the differences were not significant. E&M coding improved from a poor sensitivity rate of 21% to significantly better rate of 55%. The study also identified a poor level of accuracy concerning the capture of insurance information (kappas = .18 and .16) that conservatively indicated approximately $1.35 million dollars of missed third party collections within the past year. This study provides a model to improve the quality ADS data that may enhance the organization's ability to efficiently and effectively identify clinical process improvements, make sound resource allocation decisions, increase third party collections, and conduct outcomes research.

Table of Contents

Introduction 7

Conditions Which Prompted the Study 7

Statement of the Problem 14

Literature Review 15

Purpose 25

Methods and Procedures 25

Results 30

Discussion 34

Conclusion and Recommendations 45

APPENDIX A (ADS Encounter Form - Front) 50

APPENDIX B (ADS Encounter Form - Back) 51

APPENDIX C (ADS Data Accuracy Worksheet) 52

APPENDIX D (Evaluation and Management Coding Matrix - "DocForm") 53

APPENDIX E (ADS Encounter Form ICD-9-CM Template - Scheduled Appointments) 54

APPENDIX F (ADS Encounter Form ICD-9-CM Template - Walk-in Appointments) 55

References 56

List of Tables

Table 1 (Summary of ICD-9-CM Coding Accuracy) 61

Table 2 (Comparison Before and After Interventions of ICD-9-CM Coding Rates) 62

Table 3 (Comparison Between Medical Staff and GME students of ICD-9-CM Coding Rates).63

Table 4 (Comparison of Non-Adjusted and Adjusted ICD-9-CM Rates) 64

Table 5 (Comparison Between ADS Form and ADS Database ICD-9-CM Rates) 65

Table 6 (Summary of Evaluation and Management Coding Rates) 66

Table 7 (Comparison Before and After of Evaluation and Management Coding Rates) 67

Table 8 (Summary of Under, Over, and Inappropriate Level of Service E&M Coding) 68

Table 9 (Summary of Insurance Indicator Accuracy) 69

List of Figures

Figure 1 (WRAMC average daily occupancy from October 1995 to October 1997) 70

Figure 2 (WRAMC total clinic visits from October 1995 to October 1997) 71

Figure 3 (Percentages of marked indicators, unmarked indicators, and insured missed) 72

Introduction

Of all the activities involved in managing health care, the use of data in health care management continues to take on ever increasing importance within health care organizations. The Military Health System (MHS) is no exception. As the MHS enters competitively into the managed care environment, it is implementing new information systems to help capture the necessary data to better evaluate and manage quality, cost, and access of care. One such system, recently implemented at Walter Reed Army Medical Center (WRAMC), is called the Ambulatory Data System (ADS).

Similar to an insurance claims database, ADS captures diagnostic and procedural information at the patient-provider level for all outpatient visits. The intent of ADS is to provide military treatment facilities (MTFs) clinicians, managers, and executives a tool for business process improvements, clinical practice patterns, and outcome-based research (Hart & Conners, 1996). Although studies suggest this type of administrative data can be an appropriate source for quality management and research, there is a major underlying concern consistently addressed throughout the literature pertaining to data accuracy (Parente, Weiner, Garnick, Richards, Fowles, Lawthers, Chandler, & Palmer, 1995). Recent studies continue to raise questions about the accuracy of diagnostic and procedural coding (Fisher, Whaley, Krushat, Malenka, Fleming, Baron, & Hsia, 1992). Hence, this project addresses the issue of ADS coding accuracy and interventions to improve its accuracy.

Conditions that Prompted the Study

The MHS provides more than 50 million outpatient visits annually (USAMEDCOM, 1997). Recognizing the importance of ambulatory data collection, the Office of the Assistant Secretary of Defense, Health Affairs (OASD (HA)) issued a directive January 1995 to collect diagnoses, procedures, and insurance information for all ambulatory encounters within the MHS (Martin, 1995). To collect this data, ADS was selected as an interim solution to provide ambulatory data for TRICARE managed care support contract negotiations, to collect diagnostic and procedural data for health care determination, and to increase returns on third party collections (TPC) (Freeman, 1996).

ADS supports the MHS by collecting and reporting ambulatory encounter information that includes: (a) patient demographic information; (b) clinic specific diagnoses using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes; (b) Current Procedural Terminology (CPT) codes; (c) Evaluation and Management (E&M) codes; and third party health insurance information. It is a form-based data collection tool derived from "master bills" used in private medical practice. The actual ADS encounter form is a variation of the master bill used by the Henry Ford Clinic in Detroit (ADS Guideline, 1997) and contains space for a limited number of preprinted ICD-9-CM, CPT, and E&M codes (58, 38, and 19 codes respectively) tailored specifically by each clinic. Space is also provided on the back of the ADS encounter form for the provider to write-in three additional ICD-9-CM and three CPT codes not preprinted on the front of the form (see Appendix A - ADS Encounter Form). After an encounter, the provider marks or "bubbles-in" the appropriate primary and secondary codes on the ADS encounter form to reflect diagnoses and procedures written on the outpatient record - the Standard Form (SF) 600. The ADS encounter form is then scanned to capture all the data "bubbled-in" on the form.

ADS is an automated means to better evaluate, understand, and manage ambulatory encounters in a MHS managed care environment of decreasing resources and increasing requirements. There are five main goals of ADS for the MHS: (a) to increase accuracy in collecting and reporting workload data; (b) to achieve business process improvements through streamlined administrative procedures; (c) to provide a means for clinic managers to have a true picture of the clinic practice patterns; (d) to meet Residency Review Committee (RCC) requirements relating to encounter history of providers; and (e) to provide a baseline data set which supports outcomes based research (USAMEDCOM, 1997).

Within two years of the OASD(HA) directive, ADS has been implemented in approximately 34 U.S. Army MTFs, 35 U.S. Navy MTFs, and 83 U.S. Air Force MTFs (OASD(HA), 1997). Among the last of the MTFs to receive ADS within the MHS, WRAMC began implementation October 1996 and achieved full operational status of ADS at 62 clinic sites within the medical center by March 1997. Due to the short implementation time period and under extreme pressure from OASD(HA) to become fully operational, the medical center's training and education focused primarily on ADS compliance. Compliance is defined as the ratio of the Standard Ambulatory Data Record (SADR) to the Worldwide Workload Report (WWR). The SADR indicates the number of successfully scanned ADS encounter forms whereas the WWR indicates the number of actual visits from CHCS for a particular month. The data quality standard for this metric are 100% or greater and is the only ADS metric used by The Army Surgeon General (Burzynski, 1997). By the end of the implementation, WRAMC quickly became one of the top MTFs within the AMEDD in terms of ADS compliance and has since consistently maintained that status (Elias, 1997).

Although WRAMC has maintained among the highest overall compliance rates within the AMEDD (Elias, 1997), this metric only measures the organization's compliance in successfully scanning the encounter. It does not measure the quality of the data captured. As a result, WRAMC's Utilization Management (UM) office has raised concerns over the accuracy of

ADS data in terms of effectively and efficiently using the data to manage quality, cost, and access of ambulatory visits within the medical center (Phillips, 1997).

UM's concerns center around its ultimate goal to maximize appropriate care and minimize or eliminate inappropriate care (DoD UM Policy, 1996). While the primary focus of UM is still on inpatient activities, this is quickly changing as WRAMC transitions into the managed care environment. UM has played a major role in facilitating one of the most significant changes at WRAMC within the past two years - the shift from inpatient services to outpatient care. During this period, WRAMC's average daily occupancy has shown a downward trend from 488 beds in October 1995 to 229 beds in October 1997 (see Figure 1) while it's clinic visits have shown an upward trend from approximately 43,600 visits in October 1995 to 49,500 in October 1997 (see Figure 2). As a means of efficiently and effectively executing UM for this high volume of ambulatory care, the UM office is turning to the ADS as a potential data source for outpatient visit analyses.

There are other major conditions that have prompted a need to study ADS accuracy at WRAMC. These conditions can be grouped in the categories of quality, cost, and access of care. When assessing quality of care, the criteria or measurements that are selected must be both valid and reliable (Weitzman, 1995). In recent years, the reliability of coding medical procedures has improved on inpatient hospital data sets; but the research on the accuracy of outpatient services is limited (Garnick, Hendricks, & Comstock, 1994). Diagnoses and procedures captured by ADS represent the core of medical care utilization data; however, there have been no published studies within the MHS that address coding accuracy of the ADS data. As a result, there is no data to substantiate or refute the general perception among administrators and providers that ADS data quality is poor (Shroeder, 1997). The only formal study conducted within WRAMC regarding coding accuracy was a recent inpatient study that only reinforced the questionable accuracy of outpatient coding in the medical center.

This inpatient study focused on the coding accuracy of Diagnosis Related Groups (DRGs) and was conducted at WRAMC in February 1997. A data trend analysis of one hundred medical records was conducted to determine whether physician documentation levels, consequent data quality, and ICD-9-CM coding assignments were resulting in optimal reimbursement, accurate clinical severity, and appropriate comparative benchmarking (Phillips, 1997). The analysis identified data quality problems related to inappropriate selection of principal diagnosis, missed complications/co-morbidity conditions, and inappropriate selection of principal procedures. The coding accuracy rate was determined to be approximately 30% (Moriarity, 1997). Training for both inpatient coders and providers was identified as a key recommendation to improve accuracy for WRAMC's inpatient coding. The UM office referred to this study as an example to demonstrate the need for a similar study on outpatient coding accuracy.

However, unlike the inpatient coding, WRAMC does not have outpatient coders to abstract the medical record. At WRAMC, the clinicians or administrative assistants primarily do all outpatient coding. During the implementation of ADS, there was very little education or training provided to clinic personnel pertaining to outpatient coding for the providers (Jones, 1997). Subsequent to the implementation phase of ADS, there has been no formal follow-on training or education for providers within WRAMC. Little guidance has been provided to clinic personnel on the codes that should be preprinted on the encounter forms, how to capture the appropriate codes on the encounter forms, or how to access the data from the system. This same issue of inadequate ADS training and education of clinic personnel has also been identified as a systemic problem throughout MTFs within the MHS (Schroeder, 1997).

Data capture is another primary condition related to the management of quality of care. It was one of the primary issues identified in a final after-action report by an ADS investigative team chartered by OASD(HA) in June 1997. The team was chartered to research concerns about the current operational status and system design for ADS. As a result, there were several key data capture issues identified to include: (a) insufficient number of preprinted codes available on the form; (b) codes poorly organized on encounter form; (c) coding references not readily available to clinicians making it inconvenient and time consuming to look up the most appropriate code; and (d) coding assistance was not readily available (Shroeder, 1997). The report stated these issues hinder clinicians to accurately code the encounter (Schroeder, 1997). As the MHS moves towards the development and implementation of the computer based patient record (CPR), these issues will quickly become quality of care issues when paper based record is no longer used. Finally, the report indicated a general perception that ADS has decreased the capture of third party information.

In terms of cost of care, ADS captures information that is used for third party insurance claims. Third party insurance information is captured by ADS via a "yes" and "no" insurance indicator which front desk employees are required to mark after asking patients if they have insurance. At WRAMC, if the insurance indicator is marked "yes," the ADS encounter form is then forwarded to the medical center's third party office for billing. If the indicator is marked "no," or is unmarked, the encounter form is shredded immediately after scanning. An initial query of insurance indicators not marked during the month of October 1997 for WRAMC showed that 11,905 out of 57,947 total encounters were not marked. According to the Composite Health Care System (CHCS), 1,615 of these unmarked encounters had insurance. Assuming approximately an average of $100 collection per encounter, the unmarked encounters represent a potential revenue loss of an estimated $161,500 for that month alone.

Besides third party information, ADS data will play a key role with the upcoming enrollment based capitation (EBC) methodology to be implemented in the MHS beginning October 1998. As WRAMC transitions into the managed care environment, EBC represents one of the two major managed care initiatives being implemented within the next year. The other initiative is implementation of the TRICARE managed care support contract in June 1998. EBC and TRICARE are linked together and necessitate the need for MTF commanders to maintain accountability for all the resources used by their TRICARE Prime enrolled populations. Since most aspects of the financial accounting will be fully automated under EBC, a key factor for MTF success will depend on "complete, timely, and accurate reporting of inpatient, ambulatory, ancillary, and pharmacy transactions through CHCS and ADS" (OASD (HA)/HB&P, 1997). During the execution of EBC, "prices" will be assigned to all MTF outpatient encounter records from price tables computed by OASD(HA) for each facility based on Manpower Expense and Performance Reporting System (MEPRS) and ADS data. MEPRS and ADS records are mapped to the same MEPRS work center for two primary reasons. First, without ADS transactions, WRAMC will not receive credit for providing ambulatory services to patients who are someone else's financial responsibility. Secondly, ADS transactions will determine future prices of WRAMC's ambulatory services (OASD (HA)/HB&P, 1997). The need for accurate ADS data will become increasingly important as more resource decisions are based on this data.

Finally, in terms of access to care, accurate capture of diagnoses and procedures combined with the population-based data are essential when taking epidemiological measures to determine the health needs of the population. Knowledge of service population characteristics helps managers, at all levels, to determine not only the types and amount of services needed, but also the outcomes of the services provided (Oleske, 1995). This will become critical when WRAMC fully enters managed care with TRICARE and EBC, especially in WRAMC's GIMC where the majority of primary care providers are located. The GIMC sees more ambulatory care visits than any other clinic at WRAMC and will play a key role in TRICARE as the Family Health Center and principle Prime Site for WRAMC. As a result, the Chief, General Internal Medicine Service volunteered to participate in this project to assess the accuracy and implement interventions to improve the ADS data capture within the GIMC.

In summary, the overall management of quality, cost, and access of care calls for improved information management (Southby, 1993). Success in the managed care environment will be dependent on how well a MTF can accurately track patients, their care, and the outcomes of that care over space and time (Meyer & Krakauer, 1997). With approximately 1.1 million outpatient visits a year, accurate ADS data can only enhance WRAMC's ability to successfully compete in the managed care environment.

Statement of the Problem

As ADS data becomes more accessible from the managerial level through the executive level, the information will increasingly be used to make decisions within our health care system. In order for sound decisions to be made, decision-makers must be aware of the quality of the data being used. The lack of information concerning the accuracy of ADS data presents one major problem for decision-makers. How to improve accuracy poses another major problem. As a result, this project has two primary research questions. How accurately is the GIMC coding outpatient information on the ADS encounter forms? What recommended interventions can be implemented to improve the accuracy of the data?

Literature Review

The increased cost and utilization of ambulatory care services within the United States over the past ten years has prompted an effort on the part of payers, clinicians, and health care researchers to better understand the content of the ambulatory encounter (Golfield, 1995). This has become a significant challenge given there is currently no national system for collecting ambulatory data and evaluating the quality or the cost-effectiveness of ambulatory care (Yarnall, Michener, Broadhead, Hammond, & Tse, 1995). As a result, there are many ambulatory encounter systems (AES) currently in use or under development that capture administrative data similar to claims databases (Goldfield, 1995). Claims data have been suggested as an appropriate source for quality assessment, cost and utilization studies, medical effectiveness research, and analysis of clinical outcomes. However, current studies show the coding accuracy of diagnoses and procedures of claims data remains uncertain (Quam, Ellis, Venus, Clouse, Taylor, & Leatherman, 1993; Fisher et al., 1992).

Since there are no published studies specifically addressing the accuracy of ADS coding, the literature search on this subject focused on three main objectives: (a) to provide a brief history of diagnostic coding and review the potential uses of this type of administrative data in health care; (b) to determine an appropriate method to assess the accuracy of ADS data by a review of studies focused specifically on accuracy of administrative data; and (c) to identify interventions that could improve the accuracy of ambulatory data.

Like ADS, information in administrative databases on patient's clinical conditions are mainly in the form of diagnostic codes specified by the International Classification of Diseases, Ninth Revision, Clinical Modification, more commonly known as the ICD-9-CM. The ICD-9-CM originated at the First Statistical Congress in Brussels, Belgium, in 1853 where representatives from participating nations agreed on the need for consistent coding of causes of death worldwide. Two years later, the Congress adopted, with modifications, the general disease classification principles that involved grouping by anatomical site. Following this basic scheme, the International Statistical Institute (the successor to the Congress) produced the *Classification of Causes of Death* with the intent to revise it every ten years. This further evolved into the International Classification of Diseases (ICD) which has been maintained by the World Health Organization (WHO) since the 1940s (Israel, 1978).

The WHO convened the International Conference for the Ninth Revision in Geneva in 1975 resulting in a subsequent classification of ICD-9 that contained over 9,600 codes. However, since the 1950s, the United States had developed the concept of extending the ICD for use in hospital indexing in response to a need for a more efficient basis for storage and retrieval of diagnostic data. This created additional clinical modifications to the ICD-9 that provided more precise codes to describe the clinical picture of the patient. These clinical modifications led to the development of the ICD-9-CM which includes over 10,300 codes ranging from diseases, diagnosis, symptoms, signs, and other manifestations of diseases (Iezzoni, 1990). By 1979, the ICD-9-CM became the primary classification system for diagnoses in the United States (Jones, Castillo, Hopkins, & Aaron, 1996).

The ICD-9-CM consists of a "core" classification of three digits which is the minimum requirement for reporting mortality data to the WHO (Cimino, 1996). A fourth digit (in the first decimal place) provides an additional level of detail reflecting manifestation, and the fifth digit (second decimal place) specifies the type. (Iezonni, 1990). Decisions concerning the ICD-9-CM nomenclature in the United States are made by the ICD-9-CM Coordination-Maintenance Committee at the National Center for Health Statistics, and the Health Care Financing

Administration (HCFA).

By 1983, the entire context of the ICD-9-CM was changed in the United States when HCFA sought to control rising Medicare costs by making diagnostic and procedural data the primary determinant of hospital reimbursement for Medicare through the prospective payment system (PPS). Under this system, hospitals are paid a lump sum per hospital admission as determined by the patient's diagnosis-related group (DRG). The DRG is based on the patient's primary diagnosis, up to four additional secondary ICD-9-CM diagnoses, the presence and type of major surgery, and discharge status (Iezzoni, 1995). Since the passage of the Medicare Catastrophic Coverage Act of 1988, physicians have been required by law to submit ICD-9-CM codes when billing for services provided to Medicare beneficiaries (Jones et al., 1996).

While DRGs were effective at controlling the growth of federal outlays for acute care, the utilization and cost of ambulatory care continued to rapidly increase (Duncan & Servais, 1996). As a result, the U.S. Congress passed the Omnibus Budget and Reconciliation Act of 1989 (OBRA-89) which directed HCFA to implement a similar approach to control physician costs in the ambulatory care setting for the Medicare program. The OBRA-89 required any diagnosis or procedure provided to Medicare patients to be reported in a standard format (Horner, Paris, Purvis, & Lawler 1991). From this data, HCFA determines the medical necessity and appropriate level of reimbursement based on defined limits for a given diagnosis or procedure. In addition to the ICD-9-CM, HCFA uses two additional types of codes.

The first is the Current Procedure Terminology (CPT) code. Developed by the American Medical Association in 1966, CPT codes are used as the pre-coordinated coding scheme for ambulatory diagnostic and therapeutic procedures (Cimino, 1996). Like the DRG codes, they specify information about the codes which differentiates them based on their cost and also provides information about the type of procedure.

The second type is a sub-set of the CPT called the Evaluation and Management (E&M) code. The E&M code is a crucial factor in determining physician reimbursement. The level of the E&M code assigned formulates both resource-based relative value (RBRVS) payment rates and identifies physician practice profiles from which managed care organizations base contract negotiations (Hirshcl, 1995). Historically, the guidelines for E&M code selection have been unclear and vaguely defined. The amount of time a physician spends with a patient was considered the standard for code selection. However, in 1994, the AMA developed and defined new criteria for E&M coding. The codes are designed to classify the place of service, type of patient (new or established), referral status, level of service, and intensity of the visit (brief, intermediate, complex, etc) (Garnick et al., 1994).

Using the ICD-9-CM, CPT, and E&M codes, HCFA is actively extending PPS to the ambulatory sector to control costs and is currently in the process of phasing in ambulatory PPS over the next few years (Duncan & Servais, 1996).

Besides facilitating the management of claims reimbursement, administrative databases provide an increasingly accessible and widely used source of data for healthcare administrators and providers. Health care organizations are currently using this data to assess the quality of care, evaluate hospital utilization and practice patterns, study the appropriateness of health care costs, conduct epidemiological studies, and supplement their decision support systems (Kennedy, Stern, & Crawford, 1984; Hannah, 1995). With the rise in ambulatory care over the past decade, the attention of health care professionals is now shifting to better understand the content of the ambulatory encounter. While the primary effort still focuses on cost control, there is a growing interest within the health care community in improved measures of quality and a better understanding of clinical conditions typically treated in an outpatient setting (Goldfield, 1993). In terms of quality management, the availability of demographic, diagnostic, and procedural data allows health care professionals to follow a large sample of ambulatory patients longitudinally, examine their patterns of care, and relate these to outcomes at a low cost compared with prospective data collection (Meyer & Krakauer, 1997).

Studies of the quality of care use three primary types of data: administrative data, medical records, and surveys (Garnick, Hendricks, & Comstock, 1994). Garnick et al. (1994) defined administrative data as claims filed with payers, or records maintained by health care organizations (visits, procedures, tests, etc.). Examples of currently available administrative datasets include: the HCFA's National Claims History (NCH) covering all services provided Medicare beneficiaries; State Medicaid programs' claims databases; private insurers' claims datasets; and large managed care organizations datasets. For the MHS, ADS data collected by MTFs are transmitted to the SADR that is the primary administrative dataset for ambulatory visits within MHS.

Medical records are currently the most commonly used source of information on the quality of the process of care (Weitzman, 1996). Record reviews or chart audits have been an integral part of many quality management programs, however, problems with the medical record are especially acute in the ambulatory care setting due to fragmentation. Fragmentation simply means that information on diagnosis, procedures, and outcomes cannot be linked across settings. This occurs because medical records are typically kept locally and usually include practice specific terminology (Lohr, 1988).

Surveys are also used to obtain detailed information on ambulatory care encounters for quality management. A good example is the National Ambulatory Medical Care Survey

(NAMCS). Instituted by the Federal Government in 1973, the NAMCS takes a random sample of utilization data provided by non-Federal office-based physicians, hospital outpatient departments, and emergency departments (Schappert, 1997). The NACMS database provides some insight at the national level into the nature of ambulatory care visits, and acts as a good resource for planning health services in the ambulatory care setting or providing perspectives on national patterns of utilization (Williams & Torrens, 1993).

As a primary component of quality management, utilization management has become crucial to the success of a managed care program (Kongstvedt, 1996). One fundamental function of UM is information gathering. Administrative data potentially offers the most cost-effective means of efficiently managing five key elements of UM. The five key elements are: (a) population profile based on demographic and beneficiary categories; (b) MTF profile which includes the number of clinic visits, record completion rate, top 10 diagnoses or procedures; (c) demand management which focuses on disease management by diagnosis, provider profiling, and referral patterns; (d) disease management in terms of follow-up tracking; and (e) outcomes research (Peake, 1997).

There are many advantages of administrative data over the medical record and surveys. Individuals can be tracked over time. Large population databases make it possible to allow studies of specific providers or patient types. Researchers do not influence practice patterns through data collection (i.e., no Hawthorne effect), and research costs are much lower for the collection of primary data (Steinberg, Whittle, & Anderson, 1990). However, there are limitations of administrative data for quality management.

In ambulatory care, these limitations include the lack of clinical specificity, unknown reliability and validity of diagnoses and procedures, and lack of information on severity of illnesses (Leatherman, Peterson, Heinen, & Quam, 1991). The primary concerns about ICD-9-CM information include poorly defined clinical content, assignment of primary diagnosis, and the extent of coding secondary diagnosis (Iezzoni, 1990). However, the overriding concern is the accuracy of ICD-9-CM coding.

Studies conducted by the Institute of Medicine in the 1970s found that coding of non-clinical data such as age, gender, and dates of admission were highly accurate, but diagnoses and procedures were less reliably coded (Fisher et al., 1992). The National DRG Validation Study, published in 1988, provided the most comprehensive recent assessment of accuracy of hospital discharge data and found an overall error rate of 20.8% in DRG assignment (Hsia, Krushat, Fagan, Tebbut, & Kusserow, 1988; Fisher et al., 1992). Current studies still continue to raise the accuracy of diagnostic coding as a major problem of administrative datasets.

A review of studies focused on coding accuracy revealed a mixture of inpatient and outpatient coding studies. Theses studies covered the spectrum on data accuracy in paper-based medical records, claims databases, disease registries, survey data, and clinical trial databases (Hogan & Wagner, 1997). While the majority of studies published over the last two decades primarily focused on the validity of discharge databases created from DRG coding, more recent studies within the past five years are beginning to focus on diagnostic accuracy in the ambulatory setting. Together, these studies provide methodologies applicable for the study of ADS data accuracy.

Three types of statistics form the common method among studies reviewed for calculating accuracy of the ICD-9-CM code and address both reliability and validity. Reliability represents "the extent to which a measurement instrument has consistency over time, among various versions or applications, and within the instrument itself" (Oleske, 1995). When addressed in the studies, the common statistical technique used for reliability is the *kappa* statistic (K) which represents the extent to which agreement exists beyond that expected on the bases of chance (Maclure, 1987). Fliess (1981) proposed the following scheme for assessing the strength of agreement of the *kappa* statistic: 0 to .4, poor degree of agreement; .4 to less than .75, fair to good agreement; and above .75 excellent agreement.

Two studies on inpatient administrative accuracy reported relatively high levels of agreement between re-abstracted diagnoses and administrative data. When comparing the accuracy of Medicare data, Fisher et al. (1992) found relatively good reliability among 17 principal diagnoses with *kappas* ranging from .53 to .91. A similar study that assessed the validity of the Department of Defense's Standard Inpatient Data Record (SIDR) reported *kappas* ranging from .7 to .96 for eight diagnoses (Meyer & Krakauer, 1997). While several ambulatory data studies indicate similar levels of agreement ranging from .6 to .9, a recent study conducted by Horner, Paris, Purvis, & Lawler (1991) found a low degree of agreement between the medical record and billing form in the number of diagnoses per visit (overall *kappa* = .28) and by provider type (*kappa* ranging from .15 to .48). Horner et al. (1991) refers to this as the error of omission and offers two possible reasons for the low *kappa*: poorly organized billing forms and the lack of time for the physician to look up codes.

Similar to the ADS encounter form, hospitals submit claims to Medicare using a billing form called the UB-92 that contains space for only five diagnoses. As a result, under-reporting (or under-coding) of diagnoses was the major type of error on both the inpatient and ambulatory settings (Green & Wintfeld, 1993; Horner et al., 1991). While the primary diagnosis for the visit might be captured, remaining comorbidities are underreported either because diagnoses were not abstracted from the medical records, or because diagnoses had been cut off at five fit the UB-92 format (Green & Wintfeld, 1993). While Horner et al. (1991) used the *kappa* to address error of omission on outpatient billing forms, another study by Yarnall et al. (1995) used rates to report the error of omission. These rates were defined as the total number of diagnoses found only in the chart and not in the computer, divided by the total number of diagnoses in the chart. The rates of error of omission reported by Yarnall et al. (1995) ranged from 38% to 18%.

Validity represents "the precision to which the measure truly characterizes the phenomenon being studied" (Oleski, 1995). In terms of validity, accuracy is calculated primarily using the measures of sensitivity and positive predictive value. Within this framework, a re-abstracted record is normally considered to be the "gold standard" for defining the presence of a diagnostic condition or the performance of a particular procedure. Sensitivity (true positive rate), also referred in some studies as "completeness" or "accuracy rate," is the conditional probability that a diagnosis within the specified group was coded on the original record given that a diagnosis within the specified group was actually present on the re-abstracted record (Fisher et al., 1992; Hogan & Wagner, 1997; Yarnall et al., 1995). Reported sensitivity rates of DRG diagnoses vary from 79% (Hsai et al., 1988), 94% for myocardial infarctions (Kennedy et al., 1984), 58% to 97% (Fisher et al., 1992), and 68% to 97% (Meyer & Krakauer, 1997). The most recent studies concerning outpatient diagnosis coding report slightly lower sensitivity rates that include 69% (Horner et al., 1991), 69%, and 82% (Yarnall et al., 1995).

The positive predictive value, also commonly referred to as "correctness," is the conditional probability that a diagnoses was actually present on the re-abstracted record, given that it had been coded on the billing record (Fisher et al., 1992). In the studies reporting the positive predictive value, the range varied by diagnoses from a low of 53% to a high of 96% (Fisher et al., 1992; Nazareth, King, Haines, Rangel, & Meyers, 1993).

While knowledge of specific coding problems can be used to make corrective adjustments, awareness of error patterns may help target efforts aimed at improving accuracy of administrative data (Green & Wintfeld, 1993). This leads to the final objective of identifying interventions to improve data accuracy.

DRG coding has markedly improved over the years primarily due to adequate definition, training, monitoring, and feedback (Green & Wintfeld, 1993). However, there are few studies that compare accuracy before and after an intervention. Yarnall et al. (1995) demonstrated how outpatient coding can be simplified and accuracy improved by using a computerized dictionary of practice specific diagnoses and synonyms linked to appropriate ICD-9-CM codes. The computerized dictionary was derived from a list of all the diagnoses seen by a family practice clinic. This list was reviewed and adjusted by the clinic physicians, and in all, 717 separate diagnoses were identified and converted into a "data dictionary." This intervention significantly improved the sensitivity of diagnoses from 58% to 76% (Yarnall et al., 1995).

In other studies, Fortinsky and Gutman (1981) found that structured encounter forms significantly increased the positive predictive value of diagnosis recording relative to unstructured forms. Dambro and Weiss (1988) found that periodic monitoring of data accuracy and feedback to physicians and administrative personnel improved the sensitivity of data entry.

Finally, Kennedy et al. (1984) addressed the issue of training and education by stating that transforming descriptions of disease into numerical designation requires training. While coders are properly trained, physicians are typically inadequately informed about the process of coding. "To code accurately, a working knowledge of the medical terminology and an understanding of the characteristics, terminology, and conventions of ICD-9-CM are required" (Kennedy et al., 1984).

The history of the diagnostic and procedural coding indicates a general trend towards increasing need for such data. While the infrastructure and processes for inpatient coding for DRGs has matured within the past decade throughout the healthcare system within the United States, outpatient coding is still essentially in its introduction phase. The primary difference between inpatient and outpatient coding, especially within the MHS, is that inpatient coding is predominately done by coders, while outpatient coding is done by providers. The importance of accurate ambulatory data becomes more apparent as healthcare organizations begin to turn their attention to managed care concepts such clinical practice patterns, demand management, disease management, and outcomes research. As more ambulatory encounter systems increase within the healthcare system, it is important to take action early to improve the quality of the data so that it may be used effectively and efficiently to provide better healthcare within the managed care environment.

Purpose

The primary purpose of this project is to statistically determine the coding accuracy of ICD-9-CM codes, E&M codes, and insurance indicator codes within the GIMC at WRAMC. The secondary purpose of this project is to identify, develop, and implement interventions to serve as a model for other clinics to improve coding accuracy, and then compare the accuracy after the interventions to determine if there is a significant difference. The null hypothesis of this project is the level of ADS data accuracy was the same before and after the interventions. The alternate hypothesis stated there was a significant difference of the level of ADS data accuracy before and after the interventions.

Methods and Procedures

This project is an exploratory, qualitative study of the GIMC. It is modeled similarly from the study conducted by Horner et al. (1991) which examined the degree of accuracy of billing data in an academically affiliated family practice. The subjects of this project encompassed the medical staff of the GIMC that consisted of 10 military internal medicine staff physicians, four civilian physicians, and two nurse practitioners. Also included as subjects were 45 students enrolled in graduate medical education (interns and residents) and 7 affiliated physicians from the Uniformed Services University of the Health Sciences (USUHS) who rotate through the clinic on a weekly basis. There were a total of 68 possible health care providers to sample within the setting.

The clinic averages approximately 3,500 to 4,000 visits per month as the hospital's primary care Family Health Center that will be responsible for a population of approximately 14,000 military active duty, family member, and eligible retiree beneficiaries under TRICARE. Two simple random samples, (n1 = 99) and (n2 = 105), of ambulatory SF 600s and the corresponding completed ADS encounter forms were collected during this project.

The first sample (n1 = 99), was collected during a one week period, 15-19 December 1997. The sample was collected prospectively following a major upgrade of ADS from ADS Version 1.0 to ADS Version 2.0 to prevent biasing the second sample. The upgrade included a revision of the ADS encounter form that altered the layout of the codes and reduced the amount of space available for pre-printed codes. Version 2.0 reduced the number of preprinted ICD-9-CM codes from 60 to 58, E&M codes from 20 to 19, and CPT codes from 40 to 38. To minimize further bias on the initial sample, a mock survey form was attached to the SF 600 with instructions to the provider to complete the survey form and return it to the front desk with the SF 600. The ADS encounter forms were collected after they had been inspected for quality by the clinic head nurse per the clinic's normal process and scanned by the ADS clerk. Copies of the

SF 600 and ADS encounter forms were made and secured for analysis while the originals were returned back to the clinic.

The independent variables were the SF 600s from medical staff (which included the affiliated providers) and the GME students (interns and residents) of the Internal Medicine Service. Providers not part of or affiliated with the clinic were excluded. The dependent variables of this project were the ICD-9-CM codes and the E&M codes indicated on the ADS encounter form by the provider. Secondary focus was given one other dependent variable - the insurance indicator completed by the front desk personnel on the back of the ADS encounter form.

Nominal values were assigned based on a match between the information on the SF 600 and the ADS encounter form. An encounter was considered to be eligible for inclusion if the SF 600 contained at least one diagnosis in the chart note. Encounters with chart notes that did not include an obvious or legible assessment of a diagnosis were deemed ineligible. A total of six SF 600s were deemed ineligible for the first sample, and four for the second sample.

To determine a match, an experienced and board certified coder re-abstracted the diagnoses and procedures from the SF 600 onto a worksheet (see Appendix C - ADS Data Accuracy Worksheet). The codes re-abstracted from the SF 600s were used as the "gold standard" from which the codes on the ADS encounter were compared. The ADS encounter code and the gold standard code were considered a match if the diagnoses had the same ICD-9-CM three digit classification. After the re-abstraction, the Chief, General Internal Medicine Service then reviewed all charts, re-abstracted codes, and ADS encounter form codes to make the final determination whether the codes matched or not.

The analysis incorporated the use of the *kappa* statistic, sensitivity, and positive predictive value with 95% confidence intervals (95% CI). The *kappa* statistic was used to determine the level of agreement between the number of diagnoses indicated on the gold standard and the ADS data. Fliess' (1981) scheme for assessing the strength of agreement was used where *kappa* = 0 to .4 represented a poor level of agreement; *kappa* = .4 to .75 represented a fair to good level of agreement; and *kappa* > .75 represented an excellent level of agreement.

Sensitivity examined the degree of diagnostic accuracy or completeness codes captured on the ADS encounter form. It is defined as the rate of the total number of gold standard codes re-abstracted from the SF 600 that matched those indicated on the ADS encounter form divided by the total number of gold standard codes.

The positive predictive value (PPV) measured the correctness of codes actually captured on the ADS encounter form. It is defined as the rate of the total number of gold standard codes that matched those indicated on the ADS encounter form divided by the total number of codes on the ADS encounter form.

The statistical method used to compare the independent variables between the two samples was Chi-square analysis. The sensitivities and PPVs were analyzed at a .05 level of significance (the critical Chi square value = 3.84) using a series of 2x2 Chi-square contingency tables.

The second random sample (n = 105) was collected prospectively from 12 - 18 March 1998, three months after the first sample. The methodology used for collecting the SF 600s differed from the initial sample. Providers were asked to submit all their SF 600s each day during the collection period. The sample was pulled randomly from the records submitted throughout the collection period.

Following the collection of the first sample (n1 = 99), three main interventions were developed and implemented within the GIMC from January through March 1998. The first main intervention was the revision of the GIMC's ADS encounter form. The list of the 58 pre-printed ICD-9-CM codes and the 19 E&M codes were revised. The ICD-9-CM code revision was accomplished by gathering the samples of the 100 most common diagnoses from the WRAMC GIMC from May through August 1997 (n=3,230), DOD Wide for Internal Medicine from April 1996 to June 1997 (n = 698,132), and the 1995 National Ambulatory Medical Care Survey for Internal Medicine (n = 190,115,455). From these samples, GIMC physicians selected 58 high volume ICD-9-CM codes to be preprinted on an appointment template, and 58 for a walk-in template. The codes were then organized by organ system each template for easier identification.

The E&M codes were revised by eliminating times from the code descriptions and adding descriptions of complexity next to each numeric code. This revision was based on the ADS coding guidelines recommendation that time not be used in the description of E&M codes (ADS Guidelines, 1997).

The second main intervention was the development and execution of provider coding education within the GIMC. Three coding education classes were conducted in January primarily to the clinic medical staff to "train the trainer" so the medical staff could educate the GME students. These education classes were followed up with a document highlighting the key points from the classes. This document was distributed electronically to all GIMC medical staff and GME students via the medical center's CHCS. In addition, an E&M coding matrix developed specifically for providers, called the "DocForm" (see Appendix C - Evaluation and Management Coding Matrix - "DocForm") and was distributed to all GIMC medical staff and GME students to assist providers in properly selecting the correct E&M code.

The last intervention was the implementation of a Data Dictionary. A two page Data Dictionary published by the Family Practice Management journal called "ICD-9 Codes for

Family Practice: 1997-1998" was ultimately selected. This document contained 600 high volume diagnoses that were grouped by organ systems. The data dictionary was placed in a distinct, clear plastic document holder and posted in every exam room within the GIMC for quick accessibility as a reference for providers.

Results

The first sample (n1 =99) was analyzed from 30 different providers within the GIMC before the interventions. The sample consisted of 76 encounters seen by the medical staff and 23 seen by GME students. After the interventions, the second sample (n2 =105) was analyzed from 31 different providers within the GIMC. This sample consisted of 65 encounters seen by the medical staff and 40 seen by GME students. Of the 61 total providers analyzed from the two samples, there were a total 15 providers (10 medical staff and 5 GME students) captured in both samples. Tables 1, 6, and 9 contain summaries of the accuracy rates and the applicable 95% CI for diagnosis coding, E&M coding, and insurance indicator coding respectively.

Diagnosis coding accuracy. The first set of calculations evaluated the unadjusted results for the ICD-9-CM accuracy rates for both samples by comparing the sensitivities, PPVs, and *kappas* between the gold standard diagnoses and the ICD-9-CM codes marked on the ADS encounter form (see Table 1). The overall unadjusted sensitivity rate (completeness) increased from 60% before the interventions to 67% after the interventions. The overall unadjusted PPV rate (correctness) was slightly higher and also showed an increase from 66% to 73% before and after interventions respectively. Chi-square analysis indicated no significant difference between the rates before and after the interventions (see Table 2). Both samples indicated an overall a poor level of agreement beyond chance between the number of diagnoses on the gold standard and the ICD-9-CM codes on the ADS encounter form. However, the level of agreement doubled from *kappa* = .18 to .36 after the intervention nearing the fair to good range.

The next calculations separated out the unadjusted diagnosis data by medical staff and GME students. Both groups showed an increase in accuracy after the interventions, but only the medical staff showed a significant increase in their coding accuracy (see Table 2). Comparing the coding accuracy of each group in the first sample, the medical staff's accuracy levels were higher than the GME students, but not significant (see Table 3). There was a significant difference between the medical staff and GME students in the second sample. The medical staff's accuracy levels increased from the first sample rising from 62% to 75% sensitivity, 69% to 81% PPV, and *kappa* = .22 to .43 levels of agreement. The GME students' accuracy levels also increased from the first sample rising from 54% to 65% sensitivity, 58% to 65% PPV, and *kappa* = .04 to .23 levels of agreement. Chi-square indicated a significant difference between the medical staff and the GME students in the second sample.

The accuracy of the sample was also examined with consideration of correctly sequencing the primary diagnosis for the encounter. The data was adjusted by counting all gold standard primary diagnoses coded other than the primary ICD-9-CM code as a no match. There were a total of 18 of 146 matching diagnoses adjusted to a no match in the first sample and 8 of 125 matching diagnoses adjusted to a no match from the second sample. The resulting decrease in the accuracy rates from within each sample was not significant (see Table 4). But the adjustment did show a significant difference before and after the interventions for the overall accuracy rates and the medical staff accuracy rates (see Table 2).

The final analysis for ICD-9-CM accuracy calculated the rates based of the gold standard diagnoses matching the ICD-9-CM codes actually captured in the ADS database. These rates were then compared to the unadjusted diagnosis rates from the ADS encounter forms. Table 5 shows a significant decrease in the overall sensitivity rate from a 60% rate on the ADS encounter form to a 47% sensitivity rate on the ADS database. The PPV rate increased from 66% to 70%, but the difference was not significant. There was no significant difference between the medical staff accuracy rates and the GME student PPV rates. There was a significant difference between GME student sensitivity rates decreasing from 55% to 28%.

Evaluation and Management Coding Accuracy. The second issue examined was the E&M code sensitivity and PPV rates before and after the interventions. The first sample revealed much lower accuracy rates than the diagnosis accuracy rates. The overall E&M coding sensitivity and PPV rates for the first sample were 21% and 22% respectively (see Table 6). After the interventions, the overall rates significantly increased to a 55% sensitivity rate and a 59% PPV rate. Comparing the rates before and after the interventions showed a significant difference in the overall rates and medical staff rates (see Table 7). The medical staff's E&M coding rates were lower than the GME student's rates for the first sample, however the medical staff surpassed the GME student's rates after the interventions with a significant increase from 16% to 62% sensitivity rate and 16% to 67% PPV rate. The GME student's accuracy rates showed slight increases that were not significant (see Table 6).

The extent of over-coding, under-coding, and inappropriate coding of E&M codes was also examined. Over-coding is defined as a lower level of service documented in the SF 600 that did not support a higher level E&M code marked on the ADS encounter form. Under-coding is defined as the documentation in the SF 600 supported a higher level of service than was coded on the ADS encounter form. Inappropriate coding is defined as the inappropriate category of code used (example: established patient coded as new patient) or incomplete code (no code marked). Over-coding was the most common error committed in both samples. In the first sample, Table 8 shows the medical staff's most common error was over-coding as opposed to the GME students that under-coded more. After the interventions, the medical staff's coding improved across the board with their over-coding decreasing significantly form 45% to 23%. However, the GME students, while improving their under-coding, actually over-coded and inappropriately coded more after the interventions.

Insurance indicator coding. The final issue examined during this project was accuracy of properly coding the insurance indicator on the back of the ADS encounter form. Both samples indicated low level of agreements of correctly marking either the "yes" or "no" insurance indicator. The level of agreement decreased from *kappa* = .18 before the intervention to *kappa* = .16 after the interventions. Table 9 shows a significant increase in unmarked indicators from 33 (33%) unmarked in the first sample to 80 (76%) unmarked in the second sample. Also shown in Table 9 are the number of patients who had insurance according to CHCS but were either marked as no insurance, or unmarked. Of the 99 encounters in the first sample, 26 (26%) insured patients were missed. Of the 105 encounters in the second sample, 12 (11%) insured patients were missed.

Because of the relatively high percentages of inaccurately marked indicators and insured patients not correctly identified, a WRAMC wide sample was extracted from the ADS database encompassing the months of February 1997 through January 1998 and included in Table 9 for comparison. Throughout the medical center, the level of agreement was higher than the GIMC samples with a *kappa* = .48 which represents a fair to good level of agreement. Compared to the GIMC's rates of missed insured patients, the medical center sample of missed insured patients was lower at 6% totaling 38,399 insured encounters.

Discussion

The results showed the project's null hypothesis, which stated there would be no difference in the accuracy rates before and after the interventions, was accepted in some cases (no significant difference), and rejected in others (significant difference). The first set of results tested the hypothesis by comparing diagnosis rates before and after the interventions, between the medical staff and GME students, by non-adjusted and adjusted rates, and between the ADS database. Referring back to Table 1, the overall ICD-9-CM coding accuracy was generally poor when looking at the sensitivity and PPV rates with the kappa statistic. The ranges of the kappa statistic (kappas = .04 to .43) for this project were similar to a study conducted by Horner et al. (1991) which reported kappas ranging from .19 to .43 levels of agreement. The low accuracy rates and kappas of ICD-9-CM could be attributed to the initial ADS encounter form templates and provider education.

The clinic's original selection of preprinted ICD-9-CM codes for its ADS encounter form template did not objectively reflect the high volume codes normally seen within a General Internal Medicine Clinic. The preprinted ICD-9-CM codes listed on the original template were picked based on opinion; that is, what the providers within the GIMC thought were their high volume diagnoses based on only their experience. The providers had limited themselves to one template (58 diagnoses) for all appointment types prior to this project. Although the clinic was aware ADS supported multiple templates that could be mapped to specific appointment types, additional templates were never added. To correct this, the top 100 diagnoses from the WRAMC GIMC, DOD Wide GIMCs, and the 1995 NAMCS were reviewed by providers within the GIMC for input. Based on the consensus of the clinic providers and lessons learned from the first sample, two templates were created. One template was created primarily for scheduled appointment types to reflect ICD-9-CM codes common for follow-up appointments (see Appendix E - ICD-9-CM Scheduled Appointment ADS Encounter Form Template). The second template was created for walk-in and same day appointment types to better reflect the more acute type of ICD-9-CM codes (see Appendix F - ICD-9-CM Walk-in/Same Day Appointment ADS Encounter Form Template). The two templates had 40 common codes plus 18 unique codes. This intervention provided 76 total ICD-9-CM codes that reflected high volume diagnoses which were objectively selected from both military and civilian sources.

This revision of the ICD-9-CM templates partially explained the significant increase after the interventions with the non-adjusted medical staff rates, and also with the overall adjusted rates and adjusted medical staff rates (see Table 2). However, the non-significant increases of the overall non-adjusted rates and the GME students' rates showed that this primary intervention alone did not significantly increase ICD-9-CM coding accuracy.

The chart audit of the first sample identified some common errors contributing to the low ICD-CM-9 accuracy rates. One common error was the miscoding of V codes. V codes "identify circumstances other than disease or injury when they are the primary reason for an encounter with a provider" and are non-billable codes (ADS Guidelines, 1997). A common mistake made in the first sample was not properly coding health maintenance visits (V82.9) or prescription refills (V68.1). Instead of using these codes as the primary reason for the visit, the secondary code would be the only ICD-9-CM coded. An example was a patient seen for a prescription refill for hypertension medicine. Instead of being coded V68.1 as the primary code and hypertension (401.9) as the secondary code, only 401.9 was marked on the ADS encounter form. As a result, the diagnosis of hypertension was captured as the primary reason for the visit that subsequently created a potential for a false third party claim to be generated for a non-billable encounter.

Writing-in diagnoses on the back of the encounter form for ICD-9-CM codes not found on the preprinted list was another issue found during the chart audit. If a diagnosis was not preprinted on the form, the majority of providers would not write in the diagnosis on the back. Rather, they would select the closest ICD-9-CM code found on the front. Reasons for this behavior before the interventions could be attributed to the current metrics used to manage ADS, the absence of coding reference material that could assist providers in selecting codes not pre-printed, and the manpower requirement to identify the proper code for the written-in diagnoses on the back.

The metrics used to manage ADS within WRAMC parallels the DOD's metrics that focus on scanning compliance, and not data quality. While compliance implies the capture of quality data, there are no metrics established within WRAMC for its leadership to measure and monitor ADS coding accuracy. Establishing a metric emphasizes the importance of what is being measured and provides feedback for performance improvement. Since the current emphasis is placed on timely completion of the encounter, there is little incentive for providers to spend any more time than absolutely necessary to code the correct diagnosis. The absence of reference materials available to the provider only added an additional barrier to code accurately. Without reference materials, the providers had no means to look up codes for diagnoses they needed to write-in on the back. Because of this, some providers within the GIMC expressed a reluctance to write-in diagnoses on the back because of the additional workload it placed on the clinic's head nurse. As an additional duty, the head nurse is responsible for the quality inspections of the forms after they are completed to insure they will successfully scan. Part of the quality inspection is to insure written-in diagnoses are coded correctly. Considering the head nurse spends approximately 30 minutes to one hour a day inspecting an average of 200-300 ADS encounter forms (ranging from 6 to 18 seconds per encounter form), clinic providers were reluctant to create more work for the head nurse by having her look up additional codes.

A comparison between the gold standard diagnoses from the first sample and the ADS database showed a significant decrease in the overall sensitivity. This was partially related to codes written-in on the back of the encounter form. The database indicated the majority of codes written-in were not captured. This was primarily due to the fact that the diagnoses were not coded by the head nurse. The decrease was also partially related to the absence of the encounters in the database that indicated providers were not turning in their completed ADS encounter forms.

Implementation of the data dictionary was the intervention specifically targeted to improve the coding of diagnoses written-in. However, of the 105 encounters collected after the interventions, only seven total encounter forms had diagnoses written-in, and of the seven, only one was actually coded. Considering this along with the overall poor levels of agreement after the interventions, it appears providers were generally not using the data dictionary.

The chart audits confirmed what Kennedy et al. (1984) found in their study - that physicians are typically inadequately informed about the process of coding. Lessons learned from the first sample were incorporated into the second primary intervention - provider training and education. One shortcoming of the provider training was the time available to conduct the training. Only one hour blocks of time were allotted and only after the first training session was it realized that a minimum of two hour blocks were needed to adequately address the key issues regarding provider coding. The second shortcoming was the fact that only the medical staff from the clinic attended the training. It was the intent of the Chief, Internal Medicine Service to "train the trainer" so the medical staff could train the GME students. However, the time available to train the GME students was limited and there was not a process in place to confirm that all the GME students received the training before the second sample was taken. To compensate for this, training highlights on provider coding were distributed via electronic mail to all providers within the service.

The results after the interventions reflected significantly higher accuracy rates for those who attended the provider training compared to those who did not attend. While the GME students did not show significant improvement, the medical staff did show significant improvements after the interventions (adjusted and unadjusted) and also significantly higher accuracy rates than the GME students (see Table 3). The differences between the medical staff and GME students may also be attributed to other factors that include: (a) the GME students charting experience (i.e., not specifically identifying primary and secondary diagnoses in the assessment), (b) the small sample sizes of SF 600s collected from GME students ($n1 = 23$ and $n2 = 40$), and (c) the variability of GME students rotating through the clinic compared to the medical staff.

The last ICD-9-CM coding issue analyzed pertained to sequencing the primary diagnosis. Adjusting the match rates by enforcing the sequencing of the primary diagnosis showed an overall decrease of the accuracy rates; however, the comparison with the non-adjusted rates resulted in no significant difference. This indicates that sequencing the primary diagnosis was not a major issue within the GIMC.

While it is not reasonable to expect 100% accuracy with coded diagnoses, it is difficult to judge an acceptable range for sensitivity and PPV. Hogan et al. (1997) attempted to analyze and interpret results about rates of accuracy across studies and found it not feasible because of the variability in methods and the quality of the studies. Since there are no published benchmarks currently available, an acceptable range depends on the area of focus and potential impact. If the area of focus is third party reimbursement, then 66% sensitivity and 73% PPV to translate sub-optimization of billing and inaccurate billing. This can lead to lower collections and fraudulent billing. If the area of focus is adherence to practice guidelines for diagnoses such as asthma, hyperlipidemia, or diabetes, low accuracy rates may confound the ability to focus on variations in treatment for patients with specific conditions and adversely affect the overall tracking of the quality of care.

Provider education on ICD-9-CM coding and the use of reference material will become more important as the military transitions to the outpatient computer-based patient record (CPR). The CPR will most likely contain a balance of free text and coded entry (Overage, 1997). As a result, the codes selected in the CPR will be the only permanent record of the diagnosis on the patient. Low accuracy rates could result in missing diagnoses or inaccurately selected diagnoses leading to quality of care issues.

Although ICD-9-CM coding accuracy was originally the primary focus of this project, this quickly changed with the surprisingly low accuracy rates of E&M codes from the first sample (see Table 6). Upon further analysis of the results, three primary factors contributed to the low rates. Highlighting the importance of provider education, the first factor was the general lack of understanding among the providers concerning E&M codes. Providers simply did not understand the purpose of E&M codes, nor were they aware of three key components used to select E&M codes: history, examination, and medical decision making (AMA/HCFA, 1997). The second factor was inadequate documentation of the three key components in the medical chart. Table 8 indicates over-coding as the most frequently occurring error before and after the interventions. This was primarily due to providers not adequately documenting the key component of history. More specifically, one element of history that should be documented is any past, family, and/or social history taken during the encounter. Without this documentation, a lower E&M code must be selected. Finally, the last factor was the description of each E&M code on the ADS encounter form included time (i.e., "15 minutes"). The ADS coding guidelines (1997) clearly directs that time not be used in the description of E&M codes because it not a dominant factor. However, it was determined after the first sample that providers primarily selected their E&M codes based on the time within each description.

To offset these factors, three primary interventions were implemented. First, all references to time were removed from the E&M code descriptions on the ADS encounter forms and replaced with applicable abbreviated descriptions from the three key components. For example, E&M code 99212 description was changed from "Established, limited, 10 minutes" to "Established, Problem Focused, Straight forward." Next, the provider education intervention described earlier included E&M coding highlighting the three key components. The final intervention was the distribution of the E&M coding matrix to assist the provider in quickly and accurately identifying the correct code (see Appendix D - E&M Coding Matrix "DocForm"). Copies of this matrix were distributed to every provider via electronic mail.

The results after these interventions show an overall significant increase in both the sensitivity and PPV rates. As with the improvements discussed earlier with the ICD-CM-9 rates, the medical staff once again demonstrated significant improvement while the GME students did not. This is consistent with the fact the medical staff were the only providers privy to face to face instruction on E&M coding and use of the E&M coding matrix. Although the GME students did not show significant improvement, the fact that their accuracy rates did improve indicates that revising the E&M code descriptions had a positive influence on code selection.

The obvious impact of the low accuracy rates of E&M coding is inaccurate billing. Although this is not a current issue within the MHS, it is in the commercial sector. E&M codes are a crucial factor in determining physician reimbursement. They are also used by HCFA to identify under-coding and over-coding which may lead to reclaiming fees paid and imposing fines. This has future implications to the MHS as it implements the Medicare subvention demonstration project at six sites within the MHS. Medicare subvention enables the DoD to enroll its Medicare-eligible military retirees into the TRICARE Prime program and receive Medicare reimbursement. To do so, MTFs must meet the same terms as Health Maintenance Organizations (HMOs) serving Medicare. Thus, inaccurate E&M coding may put MTFs at full risk for future audits by HCFA.

Accurate E&M coding adds another dimension to workload reporting. Rather than simply recording a "tick" mark for each encounter, the E&M code allows for the complexity of the visit to be determined. Thus, a report showing five providers each seeing 100 patients a week can give a supervisor a better picture of their productivity by showing the complexity of the cases seen by each provider.

Accurately coding diagnoses and E&M codes directly impacts the ability of the third party collections (TPC) program at WRAMC to optimize its collections from eligible third party payers. Before the TPC office can generate a claim, the clinic must first accurately identify patients who have insurance. The results from the GIMC's insurance indicator accuracy indicate serious problems in capturing third party information from patients (see Table 9). To reduce the amount of insurance indicators either not marked or marked incorrectly, two interventions were recommended.

The first intervention involved enhancing the process of completing the insurance indicator. The original process involved the patient presenting at the front desk to check-in. Prior to this, ADS prints encounter forms for all scheduled appointments automatically printed at 0300 a.m. each morning for that day's appointments. ADS Encounter forms for walk-in/same day appointments are generated after the patient is scheduled at the front desk in CHCS. Within the GIMC, the front desk employee uses the same computer for both CHCS and ADS.

Once the patient checks-in, the front desk employee verifies patient information and insurance information that CHCS downloaded on the ADS form, writes-in and codes any corrections to this information on the ADS form, and then attaches the ADS encounter form with the SF 600 and places the forms in the licensed practical nurse's (LPN) box. The patient is asked to wait. The LPN then takes the forms from the box, calls the patient, and takes the patient's vitals. The patient returns to the waiting room, and the LPN then places the forms with the vitals in the provider's box. The provider is then the last to see the patient before the patient leaves the clinic. Through this entire process, the front desk employees are the only clinic personnel inquiring about third party insurance. This becomes difficult and information is missed when the queue to the front desk increases and the volume of telephone calls increase.

The recommended intervention was to alter this process by having the LPNs also ask the patients if they have insurance (as a back up to the front desk personnel). This would function as a verification of information already captured by the front desk personnel, and also the capture of missed information by the front desk personnel. This recommendation was accepted by the Chief, General Internal Medicine Service, but was not accepted by the clinic's head nurse. As a result, the process was not changed before the second sample was collected.

The second intervention recommended was to increase patient education materials on third party insurance within the clinic. Materials such as professionally made signs placed in prominent places, handouts, and videos explaining reasons why it is important for patients to provide insurance information. Although the clinic administrator agreed with the recommendation, no action was taken. As a result, no interventions were implemented to improve insurance indicator accuracy and therefore the percentage of unmarked indicators more than doubled from 33% in the first sample to 76% in the second sample. Compared to the medical center's overall indicator accuracy, the GIMC's indicator accuracy rated poorly (see Figure 1).

Averaging approximately $28,300 per month in collections, the GIMC maintains the highest average of outpatient third party collections among all the clinics within WRAMC and represents approximately 17.7% of the total outpatient collections. To determine the potential increase in revenue for the GIMC, an estimate was computed using the medical center's 6% insured patients missed (most conservative), an average collection of $100 per insured encounter patient, and an average of 3,500 visits per month for the GIMC. Based on the assumption of 20% of the insured missed were billable, the GIMC has the potential to increase its third party collections revenue by approximately 12% from an average of $28,300 to an average of $32,500 per month. This represents a potential increase of an additional $50,400 per year for GIMC alone. Expanding this estimate throughout the medical center, the medical center averaged approximately 58,000 visits from February 1997 to January 1998. Again, assuming that 6% of insured patients were missed, 20% of the insured missed were billable, and an average collection was $100 per encounter, the medical center is estimated to have an potential opportunity to collect an additional $69,600 per month equating to an additional $835,200 a year.

The potential for an additional $835,200 per year in outpatient collections can be considered a conservative estimate because it does not account for two important factors that could potentially increase the value. First, these estimates do not factor in the population of unmarked indictors where insurance is unknown. Considering the medical center had a total of 105,017 unmarked indicators over an eleven month period where the patient's insurance status is unknown, approximately 2% can safely be assumed to have billable insurance which represents approximately another $200,000 per year.

The second factor that the above estimates do not account for is the accuracy of the provider coding. Since the TPC office does not have outpatient coders and primarily uses the ADS encounter forms to generate bills, it is dependent on the providers to accurately code. Given the generally poor level of accuracy within the GIMC, it is clear the TPC is not optimizing the bills for the GIMC.

There were several limitations to this project that must be addressed. The first limitation was the sample size. Due to the extensive time required to collect and re-abstract the records, each sample was limited to approximately 100 records. Random sampling was used to give more statistical power to the overall sample. However, when the sample was sub-divided, only 23 records represented the GME students in the first sample resulting in a sample size not sufficiently large enough to approximate a normal probability distribution. Another limitation to this project was possible bias in the collection of the second sample. While the clinic was relatively unaware of the reason for the collection of the first sample, more were aware of the purpose of the second sample. The Hawthorne effect should also be considered as a factor for increases in the accuracy rates, especially with the medical staff. Lastly, re-abstraction and matching diagnoses were limiting factors. The accuracy rates are dependent on the level of experience of the coder and the physician reviewing the records. While both the coder and physician were highly experienced, consideration must be taken that a different coder and physician might arrive at different rates.

Conclusions and Recommendations

This project addressed two primary research questions. First, how accurately was the GIMC coding outpatient information on the ADS encounter forms? And secondly, what recommended interventions could be implemented to improve the accuracy of the data? The overall findings indicated a generally poor level of coding accuracy within the GIMC, especially in the areas of E&M coding and insurance indicator coding. This project addressed the first research question by establishing baseline metrics for ICD-9-CM codes, E&M codes, and insurance indicators. The sensitivity, PPV, and kappa values presented can be used to establish future accuracy trends within the clinic, compare accuracy rates among other GIMCs within the MHS, and give decision makers within the medical center insight into the general quality of ADS data.

The second research question was partially addressed by the interventions actually implemented during the project. However, additional possible interventions were identified and will be presented later in this section as a part of the project's overall recommendations. Among the interventions that were implemented during this project, there were two that appeared to have the most impact on coding accuracy- the revision of the ADS encounter form templates and the provider education. By taking an objective look at the high volume codes DOD wide and within the commercial sector, the results showed that some improvement could be achieved by this alone. While the method used to revise the templates can quickly and easily be implemented within any clinic, to make significant improvements in accuracy, this project showed provider education is necessary. The significant improvements by the providers who received the provider training demonstrate the value in pursuing additional coding education for the providers throughout WRAMC.

This project also reinforced the need to implement interventions directed toward improving insurance indicator marking. Based on the data available from the two clinic samples, and the medical center sample, there is potential to approximately double the third party collections for outpatient visits for the medical center. At a minimum, the results of this project merit a focused study of insurance indicators and the third party collections process.

Compared to other information systems within the medical center, ADS is a relatively new system that represents major change in the growing trend toward better managing the quality, cost, and access of outpatient services. With change comes resistance. Resistance towards ADS by both administrative and clinical staff was observed during this project. Overcoming the impediment of resistance to change is an important factor towards improving ADS coding accuracy. The following recommendations are offered as means to facilitate change and improve coding accuracy.

The first overall recommendation is to identify an internal change agent. While this project acted as an external change agent for the GIMC, there is need for an internal change agent to continue to seek ways to improve data accuracy. The same is needed within the medical center. While OSD(HA) has played the role of the external change agent to ensure compliance, the medical center needs to identify an internal change agent to ensure data quality. Since physicians are the primary source of data input, the change agent must be a physician interested and motivated to improve quality of the data collected. This primary role of this "physician champion" would be to communicate the importance of provider coding to other physicians.

A second recommendation is to actively demonstrate executive commitment towards improving ADS accuracy as essential. This can be done by strategically incorporating goals and metrics specifically directed towards ADS data quality in the medical center's strategic plan. It can also be done operationally by initiating routine updates on issues and metrics concerning ADS accuracy into meetings such the Executive Committee of the Medical and Administrative Staff (ECMAS) or the Commander's Update. This project's methodology for measuring the accuracy of ICD-9-CM, E&M, and CPT codes can be used as a model to establish ADS data quality metrics within WRAMC. By initially tracking the sensitivities, PPVs, and kappa's via random samples collected medical center wide, feedback may be provided to the executive staff and medical staff in terms of trends, i.e., whether or not ADS data quality is improving. The same concept could be applied to insurance indicators by setting a goals and tracking the percent of indicators marked on a monthly basis. An example may be setting a short-term goals to reduce the amount of indicators not marked from 19% to less than 5%, and to increase the amount of indicator marked "yes" from 9% to 15%.

Another essential recommendation is to initiate ongoing provider education and training. Side products emerged from this project that can be used throughout the medical center. One product was a two hour class on the fundamentals of provider coding. The training should be given annually to every clinical service and also to interns and GME students during their orientation prior to beginning their training programs. Besides provider coding, emphasis should also be place on the reasons why provider coding is becoming more important. Along with the training, another side product is the E&M coding matrix that should be distributed to every provider. This can be easily done electronically.

Next, WRAMC should invest in ICD-9-CM and CPT manuals. A large volume of ICD-9-CM and CPT manuals needs to be procured and made readily available to each clinic within the medical center. This will demonstrate the importance placed by the leadership on coding correctly while providing the clinician a means to look up diagnoses and procedures not preprinted on the ADS encounter form.

The Patient Administration Division should consider creating an outpatient coding section. Similar to the structure of the DRG coordinator and inpatient coders, creating positions for an outpatient coding coordinator and outpatient coders needs to be examined. This section would be the point of contact for all questions and issues concerning ADS coding and would code the medical charts of all insured ambulatory patient visits (APVs) and outpatient visits. The outpatient coders should also be responsible for taking a random sample of charts each month and assessing the accuracy of physician coding and insurance indicator marking using this project's methodology for determining sensitivities, PPVs, and kappas.

Related to the recommendation of an outpatient coding section, WRAMC should charter a process action team (PAT) for third party identification and collection. The PAT would be responsible for reviewing all the processes and procedures currently associated with third party collection; identifying innovative ways to improve the accuracy of insurance indicators; exploring the possibility of co-locating third party collections with medical records; and identifying medical center wide marketing techniques to educate the patients in the importance and benefits of providing insurance information. Marketing ideas should be actively pursued at a medical center wide level such as the use of professionally made signs, videos, and even the silent radio in the dining facility that encourages the patient to "Just Say Yes!" when asked about insurance.

Finally, this project identified 38,399 encounters of insured patients that the ADS database documents were potentially missed. Depending on the third party payer, WRAMC's third party collections program can retroactively bill two to five years back from the time of the visit. Further research into this issue is recommended by putting a team together to identify the billable encounters, pull the charts, re-abstract the encounter, and bill the third party payer.

In conclusion, this project has shown that opportunities exist to improve coding accuracy within the GIMC. It identified interventions to improve coding accuracy and methodologies to measure the sensitivities, PPVs, and kappas of ICD-9-CM, E&M, and insurance indicator coding. The results of this project demonstrate the need to initiate further assessments of ambulatory data for the entire medical center and implement the necessary interventions to improve data quality. Finally, further research and education is needed to assist the administrative and medical staff to better understand long-term implications of accurate ambulatory data in terms of third party collections, resource allocation, enrollment based capitation, demand management, disease management, practice guidelines, physician profiles, and outcomes management.

Appendix A - ADS Encounter Form (Front)

AMBULATORY ENCOUNTER SUMMARY

ICD-9-CM DIAGNOSES

477.9	Al: Allergic rhinitis
995.3	Al: Allergy unspecified
786.50	Cardio: Chest pain
401.9	Cardio: Hypertension
785.1	Cardio: Palpitations
782.1	Derm: Rash
250.00	Endo: AODM uncomplicated
272.4	Endo: Hyperlipidemia
244.9	Endo: Hypothyroidism
627.2	Endo: Menopausal states
789.00	GI: Abdominal pain
564.0	GI: Constipation
787.91	GI: Diarrhea
536.8	GI: Dyspepsia
530.81	GI: Esophageal reflux
558.9	GI: Gastroenteritis
455.6	GI: Hemorrhoids
573.3	GI: Hepatitis
305.00	Gen: Alcohol abuse
V65.49	Gen: Counseling
995.2	Gen: Drug react/problem
V70.5	Gen: Mil Physical Exam
278.00	Gen: Obesity
V68.1	Gen: Prescription refill
V82.9	Gen: Screening unsp cond
305.1	Gen: Tobacco use disorder
285.9	H/O: Anemia
466.0	ID: Acute bronchitis
462	ID: Acute pharyngitis
465.9	ID: Acute resp infection
461.9	ID: Acute sinusitis
682.9	ID: Cellulitis & abscess
372.30	ID: Conjunctivitis
382.9	ID: Otitis media
486	ID: Pneumonia
599.0	ID: UTI
616.10	ID: Vaginitis & vulvovag
079.99	ID: Viral infection
784.0	Neuro: Headache
435.9	Neuro: TIA
300.00	Psych: Anxiety state
311	Psych: Depressive disorder
493.90	Resp: Asthma
496	Resp: COPD
786.2	Resp: Cough
786.09	Resp: Shortness of breath
724.5	Rheum: Backache
727.3	Rheum: Bursitis
274.9	Rheum: Gout
719.40	Rheum: Joint pain
729.5	Rheum: Limb pain
715.00	Rheum: Osteoarthritis
848.9	Rheum: Sprains & strains
726.90	Rheum: Tendonitis
782.2	Signs: Edema
780.4	Sym: Dizziness giddiness
780.7	Sym: Malaise & fatigue
600	Uro: BPH

EVALUATION AND MANAGEMENT

99201	New, Prob focus/Str fwd
99202	New, Exp prob/Low complex
99203	New, Detailed/Mod complex
99204	New, Comprehensive
99205	New, Comp/High complexity
99211	RN/Tech only
99212	Est, Prob focus/Str fwd
99213	Est, Exp prob/Low complex
99214	Est, Detailed/Mod complex
99215	Est, Comp/High complexity
99241	Consult, Prob focus
99242	Consult, Exp Prob Focus
99243	Consult, Detailed
99244	Consult, Comprehensive
99245	Consult, High Complexity
99371	Phone consult, brief
99372	Phone consult, intermed
99373	Phone, complex/lengthy
99401	Preventive counseling,ind

CPT PROCEDURES

20605	Asp/Inj Interm Joint
20610	Asp/Inj Major Joint
93010	EKG
20550	INJ TENDON / LIGAMENT
94760	Oximetry

DISPOSITION *(Unless Inpatient)*

- Released without limitations
- Released w/work/duty limitations
- Sick at home/quarters
- Immediate referral
- Left without being seen
- Left against medical advice
- Admitted
- Expired

ADMINISTRATIVE *(Optional)*

- Consultation requested
- Referred to another provider
- Convalescent leave
- Medical board
- Medical hold

APPOINTMENT STATUS

- Appt. Sched.
- Sick-call
- Walk-in
- Tele. Consult
- Cancelled by Patient
- No-show
- Cancelled by Facility
- Inpatient
- APV

Mark here if you have address changes or corrections. Please make corrections on the back of this form.

FOR CLINIC USE ONLY

FMP/SSN:
Birthdate:

(Home)
(Work)

Insurance Company:

Tel:
Subscriber's Name:
Group:
Num:
Prov:
PCM: WALKN
BAAA/INT MED PCC WR

ID Num:
Pat Rel:
Pat SSN:
ID:

This form is subject to the Privacy Act of 1974.

APPENDIX B - ADS Encounter Form (Back)

DesignExpert™ by NCS Printed in U.S.A. Mark Reflex® EM-204940-2:6 HC02

DO NOT WRITE IN THIS SPACE

DO NOT USE PENS THAT HAVE "BLEED THROUGH" INK

OTHER DIAGNOSES — ICD Code #1 — ICD Code #2 — ICD Code #3 — DIAGNOSES

Diagnosis: ______________________
Description

Diagnosis: ______________________
Description

Diagnosis: ______________________
Description

OTHER PROCEDURES/EVALUATION & MGMT — CPT Code #1 — CPT Code #2 — CPT Code #3 — PROCEDURES

Procedure: ______________________
Description

Procedure: ______________________
Description

Procedure: ______________________
Description

NEW PRIMARY PROVIDER

PROVIDER NUMBER

INSURANCE INFORMATION

Do You Have HEALTH INSURANCE Other Than MEDICARE Or CHAMPUS? ○ Yes ○ No

Has Any Information Changed Since Your Last Visit? (If Yes, Please Make Changes) ○ Yes ○ No

Insurance Company Name: ______________________ Phone No.: ____________

Insurance Company Address ______________________

Insurance ID No.: ____________ Group Name: ____________ Group No.: ____________

Subscriber's Name: ____________ Patient's Relationship to Subscriber: ____________

ADDITIONAL PROVIDER

PROVIDER NUMBER

PROVIDER ROLE

○ Assisting Provider
○ Supervising Provider
○ Nurse
○ Para-Professional

ADDRESS CHANGES/CORRECTIONS

Name: ______________________ SSN: ____________

Address: ______________________

City: ______________________

State: ____________ Zip: ____________

Patient Cat.: ____________ Other: ____________

Home Phone No.: (______) ________ - ________

Work Phone No.: (______) ________ - ________

Sponsors Work Phone No.: (______) ________ - ________

ADDITIONAL PROVIDER

PROVIDER NUMBER

PROVIDER ROLE

○ Assisting Provider
○ Supervising Provider
○ Nurse
○ Para-Professional

FOR OFFICIAL USE ONLY

FMP	SPONSOR SOCIAL SECURITY NUMBER	TIME	DATE MO.	DATE DAY	DATE YR.

APPENDIX C - ADS Data Accuracy Worksheet

AUDIT CONDUCTED BY:
PATIENT RECORD ID (First Letter Last Name + Last Four SSN):
SEX (circle): Male Female
DATE OF BIRTH:
PATIENT CATEGORY (FMP #):
TYPE OF APPOINTMENT (circle): New or Follow-up
APPOINTMENT STATUS (circle): Appt Sched, Cancelled, Walk-in, Sick Call, Tel Con, No Show
DATE/TIME OF APPOINTMENT:
PROVIDER:

DIAGNOSIS		**ICD Code**	**DESCRIPTION From Chart (Agreed by Coder & Physician)**	**Match**	**Partial Match**	**No Match**
1	CHART			**CHART = GOLD STD**		
2	CHART			**GOLD STANDARD**		
3	CHART			**GOLD STANDARD**		
4	CHART			**GOLD STANDARD**		
U	CHART			**GOLD STANDARD**		
* Total # Additional Dx's above five in Chart:				**GOLD STANDARD**		
		ICD Code	**DESCRIPTION from ADS Sheet**			
1	ADS					
2	ADS					
3	ADS					
4	ADS					
U	ADS					
O1	ADS					
O2	ADS					
O3	ADS					
		ICD Code	**DESCRIPTION from ADS Database**			
1	DBASE					
2	DBASE					
3	DBASE					
4	DBASE					
U	DBASE					

Insurance Marked on ADS Form:	Yes	No	Unmarked			
CHCS Indicates Insurance?	Yes	No				

Eval & Mgmt	**EM Code**	**CPT**	#1 Code	#2 Code	#3 Code			
CHART		CHART				**GOLD STANDARD**		
ADS		ADS						
DBASE		DBASE						

Appendix D - Evaluation and Management Coding Matrix "DocForm"

	*OFFICE/OUTPATIENT		CONSULTATION	
	New (3)	Established (2)	Office/O.P. (3)	Confirm (3)
Minimal Service		99211		
HISTORY				
Problem Focused	99201	99212	99241	99271
Exp/Prob Focused	99202	99213	99242	99272
Detailed	99203	99214	99243	99273
Comprehensive	99204/5	99215	99244/5	99274/5
EXAMINATION				
Problem Focused	99201	99212	99241	99271
Exp/Prob Focused	99202	99213	99242	99272
Detailed	99203	99214	99243	99273
Comprehensive	99204/5	99215	99244/5	99274/5
MEDICAL DECISION MAKING				
Straight Forward	99201/2	99212	99241/2	99271/2
Low Complexity	99203	99213	99243	99273
Mod Complexity	99204	99214	99244	99274
High Complexity	99205	99215	99245	99275
TIME: Total -----------Counsel/Coordination of Care-------------------				
10 minutes	99201	99212		
15 minutes		99213	99241	
20 minutes	99202			
25 minutes		99214		
30 minutes	99203		99242	
40 minutes		99215	99243	
45 minutes	99204			
60 minutes	99205		99244	
80 minutes			99245	

* Circle the codes that best represents the encounter for HISTORY, EXAM, MEDICAL DECISION MAKING that is documented in the record. For new patients, select the lowest level code. For established, select mid level code.

HISTORY

Problem Focused: CC; brief history; exam limited to one area of body; one Dx; min data reviewed
Expanded Problem focused: CC; brief history; exam of affected area + others; Limited Dxs; Limited data reviewed; low risk of complications; expected full recovery without functional impairment.
Detailed: CC; more extensive exam; pertinent past, family, and social history; multiple number of Dxs or management options; moderate amount data reviewed; moderate risk of complications
Comprehensive: CC; history of present illness; complete single system or multisystem review; complete past, family, and social history; multiple number of Dxs; extensive data reviewed; high risk.

EXAMINATION

Problem focused: Limited exam of affected body area or organ system
Expanded Problem focused: Limited exam of a affected body area or organ system and any other symptomatic or related body area or organ system.
Detailed: Extended exam of affected body area and any other symptomatic or related body areas
Comprehensive: General Multi system exam, or complete exam of a single organ systm and other symptomatic or related body area or organ systems.

DECISION MAKING

Straight forward: Minimal number Dxs and data reviewed; minimal risk of complications/morbidity
Low Complexity: Limited Dxs and amount of data reviewed; low risk of complications/morbidity
Moderate Complexity: Multiple Dxs; Moderate data reviewed; moderate risk of complications/morbidity
High complexity: Extensive number Dxs; Extensive data reviewed; high risk of complications/morbidity

APPENDIX E - ADS Encounter Form ICD-9-CM Template - Scheduled Appointments

58	ICD-9-CM	Category	DESCRIPTION
1	477.9	Allergy	Allergic Rhinitis
1	401.9	Cardiol	Essential hypertension, NOS
1	785.1	Cardiol	Palpitations
1	786.50	Cardiol	Chest pain unspecified
1	414.00	Cardiol	Coronary Atherosclerosis of Unspecifed type
1	427.31	Cardiol	Atrial fibrillation
1	427.9	Cardiol	Cardiac dsyrhythmia, unspecified
1	428.0	Cardiol	Congestive heart failure
1	443.9	Cardiol	Peripheral vascular disease
1	782.1	Derm	Rash and other unspecified skin eruption
1	244.9	Endo	Hypothyroidism Acquired Unspecified
1	250.00	Endo	Diabetes Mellitus without mention of complications
1	272.4	Endo	Hyperlipidemia, Other and unspecified
1	627.2	Endo	Menopausal or female climacteric states
1	733.00	Endo	Osteoporosis, unspecified
1	V65.49	General	Counseling
1	278.00	General	Obesity, unspecified
1	305.00	General	Alcohol abuse, unspecified
1	305.1	General	Tabacco use disorder
1	995.2	General	Unspec adverse effect of drugs, medicinals, & barb
1	V70.5	General	Military Physical Exam
1	V82.9	General	Health Maintenance
1	079.99	General	Viral Infection Unspecified
1	562.1	General	Diverticulosis of colon without hemorrhage
1	530.81	GI	Esophageal Reflux
1	536.8	GI	Dyspepsia and other specified disorders and functi
1	564.0	GI	Constipation
1	789.00	GI	Abdominal pain, unspecified site
1	533.90	GI	PUD
1	564.1	GI	Irritable colon
1	285.9	Hem Onc	Anemia, Unspecified
1	199.1	Hem Onc	Malignant neoplasms, unspecified
1	V58.61	Hem Onc	Long-term use of anticoagulants
1	465.9	ID	Acute upper respiratory infections of unspecified
1	486	ID	Pneumonia, organism unspecified
1	599.0	ID	Urinary tract infection, site not specified
1	585	Nephro	Chronic renal failure
1	435.9	Neuro	Transient ischemic attack, unspecified
1	784.0	Neuro	Headache
1	438	Neuro	CVA, late effect, unspecified
1	357.9	Neuro	Peripheral Neuropathy Unspecified
1	300.00	Psych	Anxiety state, unspecified
1	311	Psych	Depressive Disorder
1	290.0	Psych	Senile dementia, uncomplicated
1	780.52	Psych	Insomnia
1	493.90	Resp	Asthma
1	496	Resp	Chronic airways obstruction
1	786.09	Resp	Shortness of breath
1	274.9	Rheum	Gout, unspecified
1	715.00	Rheum	Osteoarthrosis, Generalized, Involving Unspecified
1	848.9	Rheum	Sprains and strains
1	724.5	Rheum	Backache, unspecified1
1	726.90	Rheum	Tendonitis
1	782.2	Signs	Edema
1	780.4	Symptoms	Dizziness and giddiness
1	780.7	Symptoms	Malaise and fatigue
1	600	Uro	Hyperplasia of prostate
1	788.30	Uro	Incontinence

APPENDIX F - ADS Encounter Form ICD-9-CM Template - Walk-in Appointments

58	ICD-9-CM	Category	DESCRIPTION
1	477.9	Allergy	Allergic Rhinitis
1	995.3	Allergy	Allergy unspecified
1	401.9	Cardiol	Essential hypertension, NOS
1	785.1	Cardiol	Palpitations
1	786.50	Cardiol	Chest pain unspecified
1	782.1	Derm	Rash and other unspecified skin eruption
1	244.9	Endo	Hypothyroidism Acquired Unspecified
1	250.00	Endo	Diabetes Mellitus without mention of complications
1	272.4	Endo	Hyperlipidemia, Other and unspecified
1	627.2	Endo	Menopausal or female climacteric states
1	V65.49	General	Counseling
1	278.00	General	Obesity, unspecified
1	305.00	General	Alcohol abuse, unspecified
1	305.1	General	Tabacco use disorder
1	995.2	General	Unspec adverse effect of drugs, medicinals, & barb
1	V70.5	General	Military Physical Exam
1	V82.9	General	Health Maintenance
1	079.99	General	Viral Infection Unspecified
1	V68.1	General	Prescription Refill
1	530.81	GI	Esophageal Reflux
1	536.8	GI	Dyspepsia and other specified disorders and functi
1	564.0	GI	Constipation
1	789.00	GI	Abdominal pain, unspecified site
1	455.6	GI	Hemorrhoids Unspecified withou mention of compl
1	558.9	GI	Gastroenteritis and colitis, other unspecified
1	573.3	GI	Hepatitis, unspecified
1	787.91	GI	Diarrhea, NOS
1	285.9	Hem Onc	Anemia, Unspecified
1	465.9	ID	Acute upper respiratory infections of unspecified
1	486	ID	Pneumonia, organism unspecified
1	599.0	ID	Urinary tract infection, site not specified
1	112.11	ID	Candidiasis of vulva and vagina
1	372.3	ID	Conjunctivitis, unspecified
1	382.9	ID	Otitis Media Unspecified
1	461.9	ID	Acute Sinusitis, Unspecified
1	462	ID	Pharyngitis Acute
1	466.0	ID	Acute Bronchitis
1	616.1	ID	Vaginitis and vulvovaginitis, Unspec
1	682.9	ID	Cellulitis and abscess of unspecified sites
1	435.9	Neuro	Transient ischemic attack, unspecified
1	784.0	Neuro	Headache
1	300.00	Psych	Anxiety state, unspecified
1	311	Psych	Depressive Disorder
1	493.90	Resp	Asthma, Unspecified type, status asthmaticus N
1	496	Resp	Chronic airways obstruction, NEC
1	780.4	Symptoms	Dizziness and giddiness
1	786.2	Resp	Cough
1	786.09	Resp	Shortness of breath
1	274.9	Rheum	Gout, unspecified
1	715.00	Rheum	Osteoarthrosis, Generalized, Involving Unspecified
1	848.9	Rheum	Sprains and strains
1	724.5	Rheum	Backache, unspecified
1	719.4	Rheum	Pain in joint
1	727.3	Rheum	Bursitis
1	729.5	Rheum	Pain of limb
1	782.2	Signs	Edema
1	780.7	Symptoms	Malaise and fatigue
1	600	Uro	Hyperplasia of prostate

REFERENCES

Ambulatory Data System (ADS) Guidelines for Diagnosis and Procedure Coding (1997). Dated April 1997. [On-line]. Available http://199.209.10.23/adshome/guidetoc.htm

Burzynski, T. (1997). Ambulatory Data System (ADS) Metric. Information Paper dated 13 March 1997. [On-line]. Available: http://lninst2.hcssa.amedd.army.mil/TSGTop.nsf

Cimino J.J., (1996). Review paper: coding systems in health care. Methods of Information in Medicine, 35: 273-284.

Dambro, M.R. & Weiss, B.D., (1988). Assessing the quality entry in a computerized medical records system. Journal of Medical Systems, 12: 181-187.

Department of Defense (DoD) Ultilization Management Policy for the Direct Care System (Draft) (1996).

Duncan, D.G. & Servais, C.S., (1996). Preparing for the new outpatient reimbursement system. Healthcare Financial Management 50(2): 42-49.

Elias, 2LT M. (1997). Personal Interview. ADS Project Officer, Directorate of Patient Administration, Walter Reed Health Care System.

Fisher E.S., Whaley, F.S., Krushat, W.M., Malenka, D.J., Fleming, C., Baron, J.A., & Hsia, D.C., (1992). The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. American Journal of Public Health 82(2): 243-248.

Fliess, L. (1981). Statistical Methods for Rates and Proportions. (2nd ed.). New York: John Wiley & Sons.

Fortinsky, R.H. & Gutman, J.D., (1981). A two-phase study of the reliability of computerized morbidity data. Journal of Family Practice, 13: 229-235.

Freeman, W.C. (1996). Preparation of Ambulatory Data System (ADS) Encounter Forms.

Memorandum dated 29 February 1996.

Garnick DW (1994). Focus on quality: Profiling physicians' practice patterns. Journal of Ambulatory Care Management 17(3):44-75.

Garnick D.W., Hendricks, A.M., & Comstock, C.B. (1994). Measuring quality of care: fundamental information from administrative data systems. International Journal of quality in Health care, 6(2): 163-177.

Garnick, D.W., Weiner, J.P., Fowles, J., Lawthers, A.G., & Palmer, R.H. (1994). A computerized system for reviewing medical records from physicians' offices. Joint Committee Journal Quality Improvement 20(12):679-94

Goldfield, N. (1993). Ambulatory encounter systems: Implications for payment and quality. Journal of Ambulatory Care Management, 16(2): 33-49.

Green, J. & Wintfeld, N., (1993). How accurate are hospital discharge data for evaluating effectiveness of care? Medical Care, 31(8): 719-731.

Hannah, K.J.,(1995). Transforming information: data management support of health Care reorganization. Journal of the American Medical Informatics Association, 2(3): 147-155.

Hart, S.E. & Connors, R.E., (1996). A resourcing decision model for military hospitals. Military Medicine, 161: 552-559.

Hirschl, N. (1995). Eternal/mystery or essential/mastery: evaluation and management coding for physician services. Journal of AHIMA, 66 (5): 14-18.

Hsia D.C., Krushat, W.M., Fagan, A.B., Tebbut, J.A., & Kusserow, M.A., (1988). Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. New England Journal of Medicine, 318: 352-355.

Hogan, W.R. & Wagner, M.M., (1997). Accuracy of data in computer based patient records. . Journal of the American Medical Informatics Association, 4: 342-355.

Horner R.D., Paris, J.A., Purvis, J.R., & Lawler, F.H., (1991). Accuracy of patient encounter and billing information in ambulatory care. Journal of Family Practice, 33: 593-598.

Iezzoni, L.I. (1990). Using administrative data to assess the quality of hospital care. International Journal of Technology Assessment in Health Care, 6: 272-281.

Isreal, R.A., (1978). The international classification of diseases. Two hundred years of development. Public Health Reports, 93: 150-152.

Jones, LTC D.L. (1997). Personal Interview. Chief, Internal Medicine Clinic, Walter Reed Health Care System.

Jones, M.K., Castillo, L.A., Hopkins, C.A., & Aaron, W.S. (Eds.), (1996). ICD-9-CM Code Book for Physician Payment, Volumes 1 & 2, (5th ed.). Reston, Virginia: St. Anthony Publishing, Inc.

Kennedy, G.T., Stern, M.P., & Crawford, M.H., (1984). Miscoding of hospital discharges as acute myocardial infarction: implications for surveillance programs aimed at elucidating trends in coronary artery disease. American Journal of Cardiology, 53: 1000-1002.

Kongstvedt, P.R. (Ed.), (1996). The Managed Care Handbook. (3rd ed.). Gaithersburg, Maryland: Aspen Publishers, Inc.

Leatherman S., Peterson, E. Heinen, L., & Quam, L., (1991). Quality screening and management using claims data in a managed care setting. Quality Review Bulletin, 17(11): 349-359.

Lohr, K.N. (1988). Current issues in quality of care. Health Affairs, 7(1): 5-18.

Maclure, M. & Willet, W.C., (1987). Misinterpretation and misuse of the *kappa* statistic. American Journal of Epidemiology, 126: 161-169.

Martin, E.D. (1995). Ambulatory Data System Deployment. Office of Assistant Secretary of Defense, Health Affairs Memorandum dated 01 December 1995. [On-line]. Available: http://www.ha.osd.mil

Meyer, G.S. & Krakauer, H. (1997). Validity of the DoD standard inpatient data record for quality management and health services research. Unpublished manuscript.

Moriarity, C. (1997). ICD-9-CM data trend Analysis prepared for: Walter Reed Army Medical Center, Washington D.C. February 1997. St. Anthony Consulting Group.

Nazareth, I., King, M., Haines, A., Rangel, L., & Myers, S., (1993). Accuracy of diagnosis of psychosis on general practice computer system. British Medical Journal,307 (6895): 32-4.

Office of the Assistant Secretary of Defense, Health Affairs (OASD (HA)), (1997). Draft - ADS Deployment Schedule, Dated 12 December 1997. [On-line]. Available: http://cba.ha.osd.mil/ads/ adsmanualpt.htm

Office of the Assistant Secretary of Defense, Health Affairs (OASD (HA))/HB&P, (1997). Enrollment Based Capitation (EBC) Implementation Guide and ADS. [On-line]. Available: http://206.39.114.129/ebc/

Oleske, D.M. (1995). Epidemiology and the Delivery of Health Care Services. Chicago, Ill: Plenum Press.

Parente, S.T., Weiner, J.P., Garnick, D.W., Richards, T.M., Fowles, J., Lawthers, A.G., Chandler, P., & Palmer, R.H., (1995). Developing a quality improvement database using health insurance data: a guided tour with application to Medicare's National Claims History file. American Journal of Medical Quality, 10(4): 162-176.

Peake, MG J.B. (1997). Utilization Management Policy (Draft). Memorandum for

Commanders Regional Commands. U.S. Army Medical Command.

Phillips, LTC J. (1997). Personal Interview. Chief, Utilization Management, Walter Reed Health Care System.

Quam, L., Ellis, L.B., Venus, P., Clouse, J., Taylor, C.G., & Leatherman, S., (1993). Using claims data for epidemiological research. Medical Care 31(6): 498-507.

Schappert, S.M., (1997). Ambulatory care visits of physician offices, hospital outpatient departments, and emergency departments: U.S. 1995. U.S. Department of Health and Human Services, Series 13 (129).

Schroeder, D. (1997). Ambulatory Data System Tiger Team After Action Report, 2 September, 1997. [On-line]. Available: http://cba.ha.osd.mil/ads/

Southby, R.F. (1993). Military Health Care in the 21st Century. Military Medicine 158,10: 637-640.

Steinberg, E.P., Whittle, J., & Anderson, G.F., (1990). Impact of claims data research on clinical practice. International Journal of Technology Assessment of Health Care, 6: 282-287.

U.S. Army Medical Command (USAMEDCOM) (1997). Ambulatory Data System. [On-line]. Available: http://140.139.12.249/medcom

Weitzman, B.C., (1995). Improving quality of care. In A.R. Kovner (Ed), Health Care Delivery in the United States (p. 385) New York, New York: Springer Publishing Company, Inc.

Williams, S.J.,(1993). Ambulatory health care services. In Williams, S.J. & Torrens, P.R. (Eds.), Introduction to Health Services (4th ed.). Albany, NY: Delmar Publishers.

Yarnall, K.S., Michener, J.L., Broadhead, W.E., Hammond, W.E., & Tse, C.J. (1995). Computer-prompted diagnostic codes. Journal of Family Practice, 40(3): 257-262.

Table 1.

Summary of ICD-9-CM Coding Accuracy

Category	n	Sensitivity	(95% CI)	PPV	(95% CI)	kappa	(95% CI)
Non-adjusted Diagnosis Rates							
Before Interventions (n1)							
Overall	99	60%	(54%, 66%)	66%	(59%, 73%)	0.18	(.10, .26)
Medical Staff	76	62%	(43%, 67%)	69%	(60%, 78%)	0.22	(.13, .32)
GME Students	23	55%	(43%,66%)	58%	(44%,73%)	0.04	(0.0, .12)
After Interventions (n2)							
Overall	105	67%	(60%, 74%)	73%	(65%, 81%)	0.36	(.28, .46)
Medical Staff	65	74%	(65%, 83%)	80%	(70%, 91%)	0.43	(.32, .56)
GME Students	40	59%	(48%, 69%)	65%	(52%, 78%)	0.23	(.10, .36)
Adjusted Diagnosis Rates							
Before Interventions (n1)							
Overall	99	53%	(47%, 59%)	58%	(51%, 65%)	0.15	(.07, .22)
Medical Staff	76	55%	(47%, 62%)	61%	(52%, 70%)	0.19	(.10, .28)
GME Students	23	49%	(38%, 62%)	51%	(38%, 61%)	0.01	(0.0, .06)
After Interventions (n2)							
Overall	105	63%	(56%, 70%)	68%	(60%, 77%)	0.28	(.19, .37)
Medical Staff	65	70%	(61%, 79%)	76%	(65%, 87%)	0.36	(.24, .47)
GME Students	40	55%	(44%, 65%)	60%	(47%, 73%)	0.13	(.03, .24)
ADS Database Accuracy Rates							
Overall	99	47%	(41%, 53%)	71%	(60%, 80%)	0.14	(.07, .21)
Medical Staff	76	55%	(42%, 62%)	67%	(61%, 82%)	0.16	(.07, .23)
GME Students	23	28%	(5%, 18%)	70%	(20%, 93%)	0.07	(0.0, .17)

Table 2.

Comparison of Before and After Interventions of ICD-9-CM Coding Rates

Category	Before Intervention		After Intervention		
	Raw	%	Raw	%	alpha = .05
Non-adjusted ICD-9-CM					
Overall					
Sensitivity	243	60%	186	67%	NS
PPV	222	66%	170	73%	NS
Medical Staff					
Sensitivity	172	62%	98	74%	S
PPV	155	69%	90	81%	S
GME student					
Sensitivity	71	55%	88	59%	NS
PPV	67	58%	80	65%	NS
Adjusted ICD-9-CM					
Overall					
Sensitivity	243	53%	186	63%	S
PPV	222	58%	170	69%	S
Medical Staff					
Sensitivity	172	55%	98	70%	S
PPV	155	61%	90	77%	S
GME student					
Sensitivity	71	49%	88	55%	NS
PPV	67	51%	80	60%	NS

Note: S = Significant difference; NS = Not Significant

Table 3.

Comparison between Medical Staff and GME Students of Overall ICD-9-CM Rates

	Medical Staff		GME Students		
Category	Raw	%	Raw	%	alpha = .05
1st Sample (Before Interventions)					
Sensitivity	172	62%	71	54%	NS
PPV	155	69%	67	58%	NS
2nd Sample (After Interventions)					
Sensitivity	98	75%	88	59%	S
PPV	90	81%	80	65%	S

Note: S = Significant difference; NS = Not Significant

Table 4.

Comparison of Non-Adjusted and Adjusted ICD-9-CM Rates for 1st and 2nd Samples

	Non-Adjusted		Adjusted		
Category	Raw	%	Raw	%	alpha = .05
1st Sample (n1 = 99)					
Overall Sensitivity	243	60%	243	53%	NS
Overall PPV	222	66%	222	58%	NS
Medical Staff Sensitivity	172	62%	172	55%	NS
Medical Staff PPV	155	69%	155	61%	NS
GME student Sensitivity	71	55%	71	49%	NS
GME student PPV	67	58%	67	51%	NS
2nd Sample (n1 = 99)					
Overall Sensitivity	186	67%	186	63%	NS
Overall PPV	170	74%	170	69%	NS
Medical Staff Sensitivity	98	74%	98	70%	NS
Medical Staff PPV	90	81%	90	77%	NS
GME student Sensitivity	88	59%	88	55%	NS
GME student PPV	80	65%	80	60%	NS

Note: S = Significant difference; NS = Not Significant

Table 5.

Comparison of Between ADS Form and ADS Database ICD-9-CM Rates

Category	Non-Adjusted ADS Form (n1) Raw	%	ADS Database Raw	%	alpha = .05
Overall					
Sensitivity	243	60%	243	47%	S
PPV	222	66%	162	70%	NS
Medical Staff					
Sensitivity	172	62%	172	55%	NS
PPV	155	69%	132	71%	NS
GME student					
Sensitivity	71	55%	71	28%	S
PPV	67	58%	30	67%	NS

Note: S = Significant difference; NS = Not Significant

Table 6.

Summary of Evaluation and Management Coding Rates

Category	n	Sensitivity	(95% CI)	PPV	(95% CI)
Before Intervention (n1)					
Overall	99	21%	(13%,29%)	22%	(13%, 31%)
Medical Staff	76	16%	(8%, 24%)	16%	(8%, 25%)
GME Students	23	39%	(19%, 59%)	41%	(17%, 64%)
After Interventions (n2)					
Overall	105	55%	(46%, 65%)	59%	(48%, 70%)
Medical Staff	65	62%	(50%, 73%)	67%	(52%, 81%)
GME Students	40	45%	(30%, 60%)	47%	(29%, 66%)
ADS Database (n1)					
Overall	99	18%	(11%, 26%)	22%	(13%, 32%)
Medical Staff	76	16%	(8%, 24%)	18%	(8%, 28%)
GME Students	23	26%	(8%, 44%)	40%	(9%, 70%)

Table 7.

Comparison: Before and After of Evaluation and Management Coding Rates

	Before Intervention		After Intervention		
Category	Raw	%	Raw	%	alpha = .05
Overall					
Sensitivity	99	21%	105	55%	S
PPV	95	22%	98	59%	S
Medical Staff					
Sensitivity	76	16%	65	62%	S
PPV	73	16%	60	67%	S
GME student					
Sensitivity	23	39%	40	45%	NS
PPV	22	41%	38	47%	NS

Note: S = Significant difference; NS = Not Significant

Table 8.

Summary of Under, Over, and Inappropriate Level of Service of E&M Coding

	Before Intervention		After Intervention	
Category	Raw #	%	Raw #	%
Overall	99	100%	105	100%
Properly coded	21	21%	58	55%
Over-coded	37	37%	26	25%
Under-coded	19	19%	7	7%
Inappropriate	22	22%	14	13%
Staff	76	100%	65	100%
Properly coded	12	16%	40	62%
Over-coded	34	45%	15	23%
Under-coded	11	14%	3	5%
Inappropriate	19	25%	7	11%
GME Students	23	100%	40	100%
Properly coded	9	39%	18	45%
Over-coded	3	13%	11	28%
Under-coded	8	35%	4	10%
Inappropriate	3	13%	7	18%

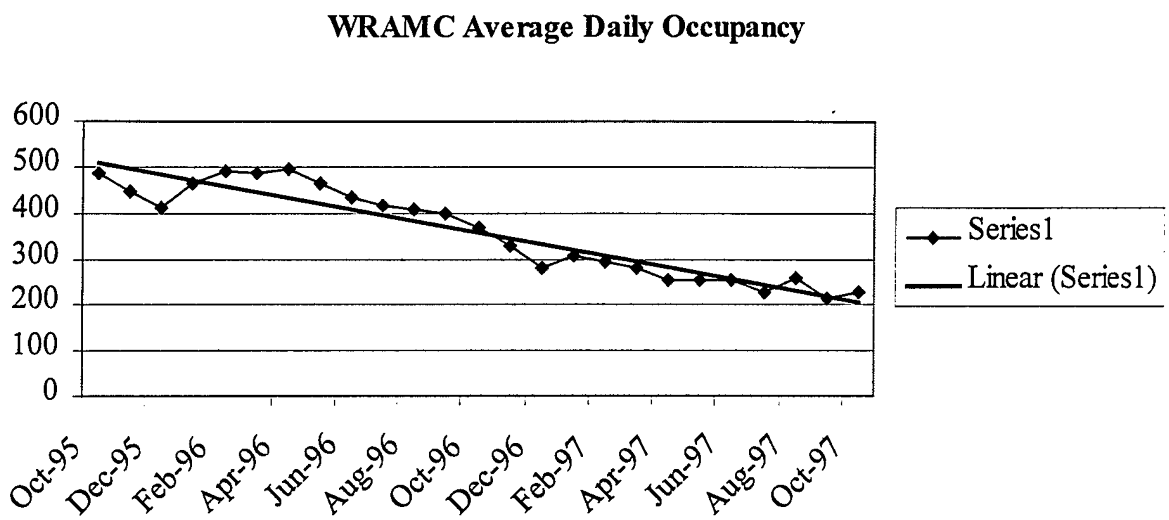
Note: Properly coded: Documentation supported E&M code
Under-coded: Documentation supported higher level of service
Over-coded: Documentation did not support level of service
Inappropriate: Inappropriate category of code used or Incomplete

Table 9.

Summary of Insurance Indicator Accuracy

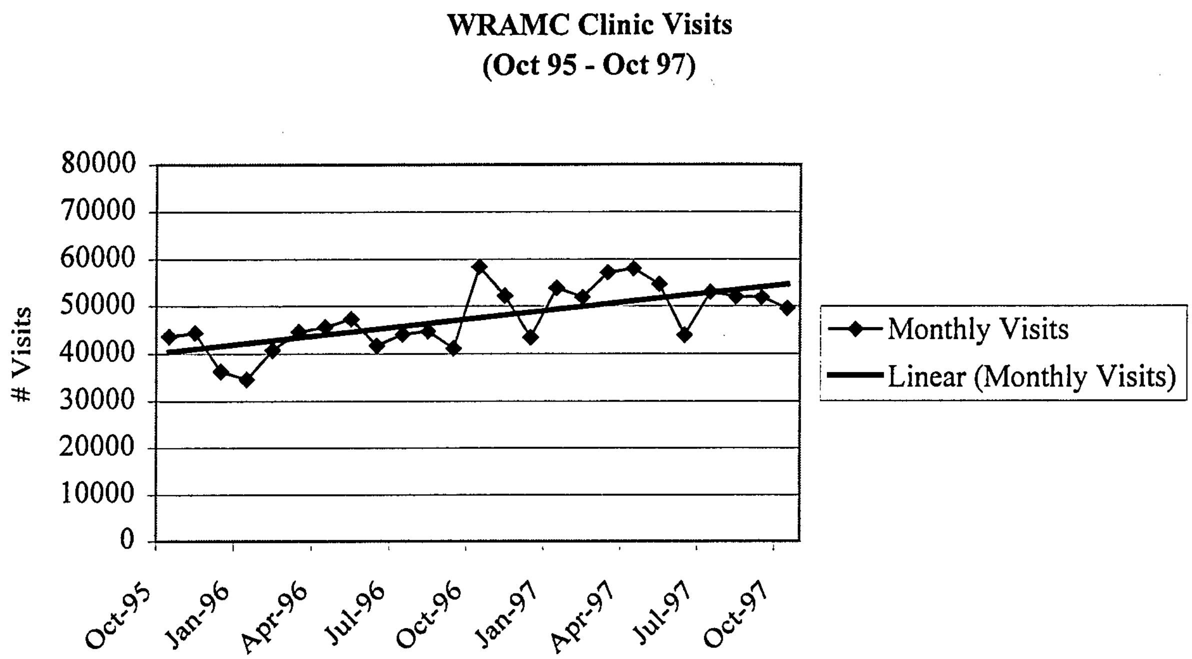
Category	Before Intervention n1	%	After Intervention n2	%	WRAH Feb 97 - Jan 98 n3	%
Sample Size	99	-	105	-	638,377	-
kappa	0.18	-	0.16	-	0.48	-
(95% CI)	(.10, .25)	-	(.11, .21)	-	(.47, .48)	-
Indicators Marked "Yes"	4	4%	4	4%	57,614	9%
CHCS Disagrees (data not in CHCS)	0	0%	0	0%	5,658	1%
Indicators Marked "No"	62	63%	21	20%	45,8611	72%
CHCS Disagrees (data indicates insurance)	12	12%	1	1%	21,265	3%
Indicators Un-marked	33	33%	80	76%	122,151	19%
CHCS Indicates insurance	14	14%	11	10%	17,134	3%
Total insured patients missed	26	26%	12	11%	38,399	6%
Total unmarked + insured missed	45	45%	81	77%	143,416	22%

Figure 1. WRAMC's average daily occupancy from October 1995 to October 1997. Average daily occupancy is defined as total bed days divided by the number of days in the period.



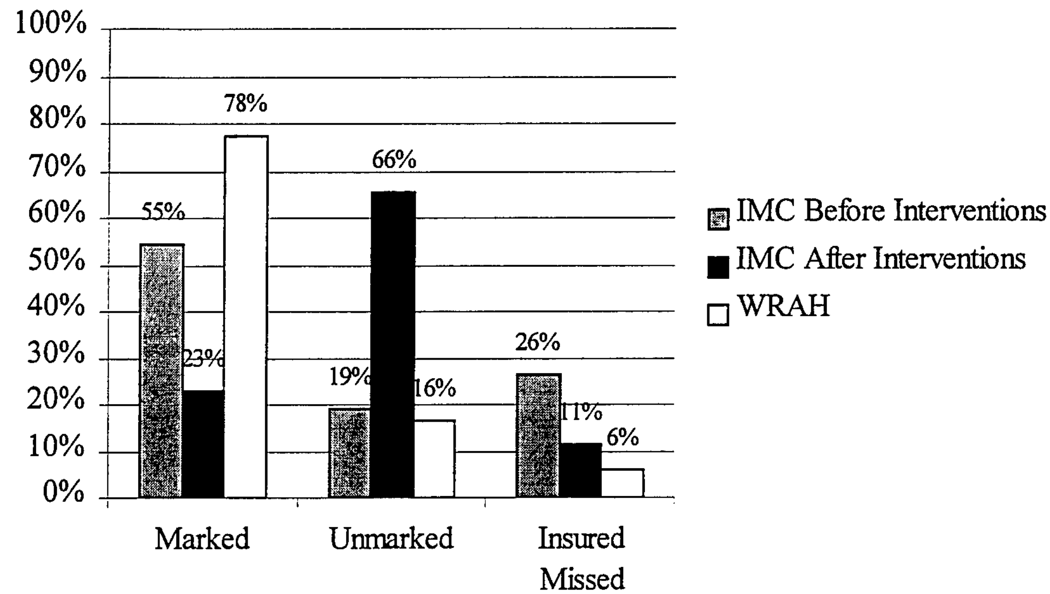
Source: CHCS

Figure 2. WRAMC's total monthly clinic visits from October 1995 to October 1997.



Source: CHCS and Clinic Reports

Figure 3. Percentages of marked indicators ("yes" and "no"), unmarked indicators (insurance unknown), and insured missed (includes unmarked plus "no" marked that CHCS indicates as having insurance). Sample sizes were GIMC before interventions (n1 = 99), GIMC after interventions (n2 = 105), and WRAH (n3 = 638,376).



Source: ADS

Table 7.

Comparison: Before and After of Evaluation and Management Coding Rates

	Before Intervention		After Intervention		
Category	Raw	%	Raw	%	alpha = .05
Overall					
Sensitivity	99	21%	105	55%	S
PPV	95	22%	98	59%	S
Medical Staff					
Sensitivity	76	16%	65	62%	S
PPV	73	16%	60	67%	S
GME student					
Sensitivity	23	39%	40	45%	NS
PPV	22	41%	38	47%	NS

Note: S = Significant difference; NS = Not Significant